The Annals of Applied Probability
2018, Vol. 28, No. 3, 1821–1855
https://doi.org/10.1214/17-AAP1344

LARGE DEVIATIONS THEORY FOR MARKOV JUMP MODELS OF CHEMICAL REACTION NETWORKS

BY ANDREA AGAZZI[*,†,1,2], AMIR DEMBO[*,3] AND JEAN-PIERRE ECKMANN[†,1]

Stanford University[*] and University of Geneva[†]

We prove a sample path Large Deviation Principle (LDP) for a class of jump processes whose rates are not uniformly Lipschitz continuous in phase space. Building on it, we further establish the corresponding Wentzell–Freidlin (W-F) (infinite time horizon) asymptotic theory. These results apply to jump Markov processes that model the dynamics of chemical reaction networks under mass action kinetics, on a microscopic scale. We provide natural sufficient topological conditions for the applicability of our LDP and W-F results. This then justifies the computation of nonequilibrium potential and exponential transition time estimates between different attractors in the large volume limit, for systems that are beyond the reach of standard chemical reaction network theory.

1. Introduction. The dynamics of chemical reactions are usually modeled by mass-action equations: A system of a polynomial ordinary differential equations which relate the evolution of concentrations of chemical compounds. These systems of equations inherit their structure from the topology of the Chemical Reaction Network (CRN) they model, and the interplay between topology and dynamics of mass action systems is the object of study of chemical reaction network theory [1, 12, 20]. These sets of ODEs approximate the interactions of the individual molecules involved. The discrete nature of chemical reaction systems can be captured by discrete models where the state of the system is given by the number of molecules of each type that are present in the reactor. In this framework, when a reaction occurs, the input molecules combine to form the output ones, and the system jumps to a new state. The dynamics of such systems are in general modeled stochastically as a pure jump Markov process [11], Example C, Section 11, whose jump rates are approximations of the reaction rates found in deterministic mass action models. Finally, assuming that the system has volume v, one can study how the stochastic dynamics of the process X_t^v describing the concentration of the

Received January 2017; revised August 2017.
[1]Supported in part by ERC advanced grant 290843 and SwissMAP.
[2]Supported in part by SNSF grant 161866.
[3]Supported in part by NSF grant DMS-1613091.
MSC2010 subject classifications. Primary 60F10, 80A30; secondary 37B25, 60J75.
Key words and phrases. Large deviation principle, Wentzell–Freidlin theory, jump Markov processes, chemical reaction networks, Lyapunov functions, toric jets.

different chemical species at time t scale with the parameter v. This is the object of study of this paper.

Similar discrete stochastic mass action kinetics models have been applied to disease propagation dynamics [27], genetic algorithms [24], and for the simulation of noisy biochemical reaction networks through the application of the so-called Gillespie algorithm [17]. Asymptotics such as limit theorems on the convergence of the stochastic trajectories toward the deterministic ones have been proven in the probability literature [11]. More recently, results on product-form steady state distributions for a certain class of CRNs have been obtained in [2, 6] and conditions for the irreducibility and ergodicity of the stochastic chemical dynamics of reaction networks have been presented in [19, 25]. Our work extends these results to the domain of large deviations theory, identifying a large class of CRNs to which that theory applies. We prove in particular that Wentzell–Freidlin exit time estimates can be applied to such systems, rigorously justifying the widespread use of potential theory [15, 16, 22] and ultimately allowing for the analysis of events that play a key role in, for example, theoretical biochemistry [3, 4] and that are not covered by deterministic mass action models, because deterministic models do not allow for transitions between different attractors.

1.1. *The model and its sample path LDP.* We consider a set of *chemical species* $\mathcal{S} = \{s_1, s_2, \dots, s_d\}$, whose interactions are described by a finite set of *reactions* $\mathcal{R} = \{r_1, r_2, \dots, r_m\}$. Throughout, we denote by $\mathbb{N}_0$ the set of natural numbers including 0. Each reaction is uniquely identified by its *substrates* (input species) and *products* (output species), and we express such a reaction as $r = \{c^r_{\text{in}} \rightharpoonup c^r_{\text{out}}\}$, with $c^r_{\text{out}}, c^r_{\text{in}} \in \mathbb{N}_0^d$ representing the multiplicity of the species $s_i \in \mathcal{S}$ in the input or output of the reaction. The set $\mathcal{C}$ of *complexes* consists of all $c^r_{\#}$ (with # = “in” or “out”), and for each reaction $r \in \mathcal{R}$ we define the *reaction vector* $c^r := c^r_{\text{out}} - c^r_{\text{in}} \in \mathbb{Z}^d$. A CRN is thus defined by the triple $(\mathcal{S}, \mathcal{C}, \mathcal{R})$.

EXAMPLE 1.1. The system

$$\varnothing \overset{r_1}{\rightharpoonup} A + B \overset{r_2}{\rightharpoonup} 2B \overset{r_3}{\rightharpoonup} A \tag{1.1}$$

is a CRN with $\mathcal{S} = \{A, B\}$ and $\mathcal{R} = \{r_1, r_2, r_3\}$. The set of complexes of these reactions is $\mathcal{C} = \{\varnothing, \{A + B\}, \{2B\}, \{A\}\} = \{(0, 0), (1, 1), (0, 2), (1, 0)\}$ [in the basis spanned by (A, B)].

In this paper, we study the behavior as a function of v of the scaled process

$$(X^v_t)_i := v^{-1}(N_t)_i, \qquad i \in 1, \dots, d,$$

where $N_t \in \mathbb{N}_0^d$ represents the number of molecules of the d species and $X^v_t \in (v^{-1}\mathbb{N}_0)^d$ denotes their number density (in mols) at time t. The interactions among molecules are then described by each reaction $r \in \mathcal{R}$ standing for a possible jump of the process $X_t^v \to X_t^v + v^{-1}c^r$, with c^r the reaction (or jump) vector associated with $r \in \mathcal{R}$. Correspondingly, X_t^v is a continuous time pure jump Markov process with generator

$$(\mathcal{L}_v f)(x) := v \sum_{r \in \mathcal{R}} \Lambda_r^{(v)}(x)\big(f\big(x+v^{-1}c^r\big) - f(x)\big) \tag{1.2}$$

for $f : (v^{-1}\mathbb{N}_0)^d \to \mathbb{R}$ and the volume-normalized mass action kinetics jump rates

$$\Lambda_r^{(v)}(x) = k_r v^{-\|c_{\text{in}}^r\|_1} \prod_{i=1}^d \binom{v x_i}{(c_{\text{in}}^r)_i} (c_{\text{in}}^r)_i! \tag{1.3}$$

for some reaction (rate) constants $k_r > 0$, where $\binom{a}{b}$ denotes the binomial coefficient which by convention is zero when $b \notin [0, a]$ and $\| \cdot \|_1$ denotes the ℓ_1-norm. The mean-field character of this model reflects the underlying assumption of homogeneous stirring of the reactor. The scaling in v of the rate constants makes them asymptotically extensive quantities in (1.2) and takes into account that it is harder for molecules to meet as v increases.

REMARK 1.2. For a fixed volume v and initial condition $X_0^v = x_0^v \in (v^{-1}\mathbb{N}_0)^d$, the process X_t^v is confined to $S_{x_0^v}^v := \{x_0^v + \{\sum_{r \in \mathcal{R}} \alpha_r c^r : \alpha \in (v^{-1}\mathbb{N}_0)^m\}\} \cap \mathbb{R}_+^d$, where $\mathbb{R}_+$ represents the set of nonnegative real numbers. Indeed, X_t^v cannot jump outside of $(v^{-1}\mathbb{N}_0)^d$ since $\Lambda_r^{(v)}(x) = 0$ for any $r \in \mathcal{R}$ such that $x + v^{-1}c^r \notin (v^{-1}\mathbb{N}_0)^d$ so the corresponding summand in (1.2) is then zero [regardless of $f(\cdot)$].

REMARK 1.3. In the limit $v \to \infty$, the sample paths of the processes X_t^v starting at $X_0^v = x_0^v \to x_0 \in \mathbb{R}_+^d$ almost surely converge—uniformly over $[0, T]$ for any $T > 0$—to the solution $\zeta(t)$ of the deterministic ODE

$$\frac{\mathrm{d}\zeta}{\mathrm{d}t} = \sum_{r \in \mathcal{R}} \lambda_r(\zeta) c^r, \qquad \zeta(0) = x_0, \tag{1.4}$$

having the asymptotic reaction rates

$$\lambda_r(x) := k_r \prod_{i=1}^d x_i^{(c_{\text{in}}^r)_i} \tag{1.5}$$

provided that a solution of (1.4) exists up to time T (see [11], Theorem 2.1, Section 11, where such a functional law of large numbers (FLLN) is derived for certain CRNs).

We show in Section 2 that under the following mild assumption on the generator $\mathcal{L}_v$ of the scaled process, the solution X_t^v to the corresponding martingale problem satisfies a sample path LDP in the supremum norm, with an explicit rate function (see Theorem 1.6). While proving this LDP we also verify that in this setting the ODE (1.4) admits global solutions (and that the FLLN of Remark 1.3 holds).

ASSUMPTION A.1. Let X_t^v be the solution of the martingale problem generated by the generator $\mathcal{L}_v$ of (1.2). We assume:

(a) There exist $b < \infty$ and a continuous, positive function $U(x)$ of compact level sets, such that for some nondecreasing function $v' : \mathbb{R}_+ \to \mathbb{R}_+$,

$$(\mathcal{L}_v U^v)(x) \le e^{bv} \qquad \forall v > v'(\|x\|_1), x \in (v^{-1}\mathbb{N}_0)^d, \tag{1.6}$$

where $U^v(\cdot)$ denotes the vth power of $U(\cdot)$.

(b) With positive probability, starting at $X_0^v = 0$ the Markov process X_t^v reaches in finite time some state x_+ in the strictly positive orthant $(v^{-1}\mathbb{N})^d$.

REMARK 1.4. The existence of a solution X_t^v to the martingale problem generated by $\mathcal{L}_v$ with initial condition $x_0^v \in (v^{-1}\mathbb{N}_0)^d$ is guaranteed by standard theory (see [26], Theorem 8.3), up to the possibility of explosion. In Lemma 2.1 we show that this possibility is ruled out by Assumption A.1.

REMARK 1.5. Assumption A.1(b) requires that all chemical species can be created, at least indirectly, starting from zero, hence from any other possible state of the system. In particular, there must exist at least one chemical reaction without substrates, namely, with $c_{\text{in}}^r = 0$. Such constant rate reactions are used, for example, in mass action models of cellular dynamics [2] and continuous-flow stirred-tank chemical reactors [12], to model inflow of chemicals from the environment (correspondingly, these CRNs often also have certain products exit the network, reflected by a mass loss in some reactions). It is possible to have an LDP without Assumption A.1(b), but then even when starting at $x_0^v \to x_0$ which is strictly positive, we may have a path of finite rate that leads to $\partial\mathbb{R}_+^d$ and stays there forever. This would create problems establishing the Wentzell–Freidlin estimates.

Proceeding to state our sample path LDP, hereafter $D_{0,T}(\mathbb{R}_+^d)$ denotes the space of *càdlàg* functions $z : [0, T] \to \mathbb{R}_+^d$ equipped with the topology of uniform convergence. For $z(\cdot)$ in the subspace $AC_{0,T}(\mathbb{R}_+^d)$ of absolutely continuous functions from $[0, T]$ to $\mathbb{R}_+^d$, let $z'(\cdot)$ denote its Radon–Nikodym derivative with respect to Lebesgue measure. Further, for $\lambda = (\lambda_r) \in \mathbb{R}_+^m$, $q = (q_r) \in \mathbb{R}_+^m$, $\xi \in \mathbb{R}^d$ and $c^r \in \mathbb{R}^d$, let

$$\begin{aligned} L(\lambda, \xi) &:= \sup_{\theta \in \mathbb{R}^d} \Big\{ \langle \theta, \xi \rangle - \sum_{r \in \mathcal{R}} \lambda_r \big[\exp(\langle \theta, c^r \rangle) - 1 \big] \Big\} \\ &= \inf \Big\{ \sum_{r \in \mathcal{R}} \Big[\lambda_r - q_r + q_r \log \frac{q_r}{\lambda_r} \Big] : q \in Q_{\mathcal{R}}(\xi) \Big\}, \end{aligned} \tag{1.7}$$

where $Q_{\mathcal{R}}(\xi) := \{q \in \mathbb{R}_+^m : \sum_{r \in \mathcal{R}} q_r c^r = \xi\}$ and $\langle \theta, \xi \rangle$ is the inner product of $\theta, \xi \in \mathbb{R}^d$.

THEOREM 1.6. *For $\lambda_r(\cdot)$ of (1.5) and any $x_0^v \to x_0 \in \mathbb{R}_+^d$, under Assumption A.1 the sample paths $\{X_t^v : t \in [0,T]\}$ with $X_0^v = x_0^v$, satisfy the LDP in $D_{0,T}(\mathbb{R}_+^d)$ with rate v and the good rate function*

$$(1.8)\quad I_{x_0,T}(z) := \begin{cases} \int_0^T L(\lambda(z(t)), z'(t))\mathrm{d}t & \text{if } z(0) = x_0 \;\&\; z \in AC_{0,T}(\mathbb{R}_+^d), \\ \infty & \text{otherwise}. \end{cases}$$

That is, for any set $\Gamma \subset D_{0,T}(\mathbb{R}_+^d)$, denoting by Γ^o and $\bar{\Gamma}$ the interior and, respectively, the closure of Γ, we have

$$(1.9)\qquad \limsup_{v\to\infty} \frac{1}{v} \log \mathbb{P}_{x_0^v}[X_t^v \in \bar{\Gamma}] \le - \inf_{z\in\bar{\Gamma}} I_{x_0,T}(z),$$

$$(1.10)\qquad \liminf_{v\to\infty} \frac{1}{v} \log \mathbb{P}_{x_0^v}[X_t^v \in \Gamma^o] \ge - \inf_{z\in\Gamma^o} I_{x_0,T}(z).$$

REMARK 1.7. The identity (1.7) is well known (see [28], Theorem 5.26), and since the function $[b - u + u\log(u/b)]$ is positive whenever $u \neq b$, it yields that the Lagrangian $L(\lambda, \xi)$ vanishes iff $\xi = \sum_{r\in\mathcal{R}} \lambda_r c^r$. Thus, the rate $I_{x_0,T}(z)$ of (1.8) is zero iff $z(\cdot)$ solves on $[0,T]$, the ODE (1.4) starting at $z(0) = x_0$ (see [28], Exercise 5.14).

1.2. *Topological stability and strongly endotactic networks.* Standard large deviations theory is not directly applicable for proving Theorem 1.6, because we need to deal with jump rates that are neither bounded away from zero, nor globally Lipschitz continuous. The diminishing jump rates at the boundary are handled by adapting our system to the framework of mean-field interacting particle systems, and thereby applying [9], Theorem 3.9, whereas Lemma 2.1 takes care of the lack of global Lipschitz continuity by employing Lyapunov stability theory to establish exponential tightness. In doing so, a most important challenge is to phrase a stability condition strong enough for such exponential tightness, and a sufficient condition for escape from the boundary (in extension of [29]), that are both applicable to a broad collection of CRNs.

This is precisely what we do next, with our topological conditions summarized by Assumption A.2 below. Specifically, given a finite set $Q \subset \mathbb{R}^d$ and a vector $w \in \mathbb{R}^d$, we call

$$Q_w := \{c \in Q : \langle w, c - c'\rangle \ge 0 \text{ for all } c' \in Q\},$$

the w-maximal subset of Q and consider the following collection of CRNs.

DEFINITION 1.8 ([18]). The network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is called *strongly endotactic* if for any nonzero $w \in \mathbb{R}^d$, the set $\mathcal{R}_w \subseteq \mathcal{R}$ of reactions such that $c_{\mathrm{in}}^r \in (\mathcal{C}_{\mathrm{in}})_w$ contains at least one reaction satisfying $\langle w, c^r\rangle < 0$ and no reaction with $\langle w, c^r\rangle > 0$.

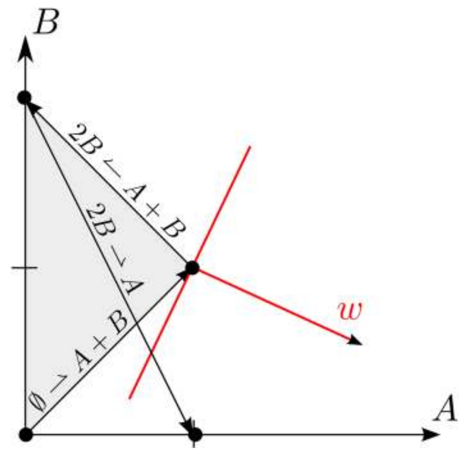

FIG. 1. *The $\mathbb{R}_+^d$-diagram for Example 1.1. A vector w in the space of complexes and the corresponding orthogonal hyperplane has been drawn in red to identify the w-maximal subset of the input complexes $(\mathcal{C}_{\text{in}})_w$: the complex $A + B$.*

This class of CRNs is well known (see [18]), and algorithms to determine if a network is strongly endotactic are devised in [21] (using variants of the simplex algorithm).

EXAMPLE 1.1 (continued). The network of Example 1.1 is represented as in Figure 1, where we identify $(\mathcal{C}_{\text{in}})_w$ by sweeping $\mathbb{R}_+^d$ with a hyperplane orthogonal to $w \in \mathbb{R}^d$ (here for $d = 2$, drawn in red), and taking the last point of $\mathcal{C}_{\text{in}}$ that such hyperplane intersected. It is easy to see that our specific network satisfies the requirements of Definition 1.8 and is therefore strongly endotactic.

While in a strongly endotactic reaction network, all reactions "point inward" with respect to the faces of the convex hull of $\mathcal{C}_{\text{in}}$ (etymologically *endo-tactic*: inward-arranged), our LDP requires addressing the following additional boundary concept.

DEFINITION 1.9. A nonempty subset $\mathcal{P} \subseteq \mathcal{S}$ is called a *siphon* if every reaction $r \in \mathcal{R}$ with at least one output from $\mathcal{P}$ also has some input species from $\mathcal{P}$.

EXAMPLE 1.10. It is readily checked that the sets $\mathcal{P} = \{A\}, \{A, B\}$ are siphons of the network

$$A \rightharpoonup 2A \rightharpoonup 3A + 2B \rightharpoonup A,$$

whereas $\{B\}$ is not.

We make the following assumption on the topological structure of CRNs. We call $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ an *asiphonic strongly endotactic* (ASE) network if it satisfies

ASSUMPTION A.2. The CRN $(\mathcal{S}, \mathcal{C}, \mathcal{R})$has the properties:

(a) It is strongly endotactic, as in Definition 1.8,
(b) It has no siphon $\mathcal{P} \subseteq \mathcal{S}$.

REMARK 1.11. Assumption A.2(b) is equivalent to finding, for any nonempty $\mathcal{P} \subseteq \mathcal{S}$, some reaction from $\mathcal{R}$ that produces at least one output in $\mathcal{P}$ while requiring no input species from $\mathcal{P}$. When this holds, then, for any state x on the $\mathcal{P}$-boundary of $\mathbb{R}_+^d$ (namely, having $x_i = 0$ for all $s_i \in \mathcal{P}$), there is some reaction of a nonvanishing rate that brings the system back to a higher-dimensional subspace of $\mathbb{R}_+^d$. Following a sequence of such jumps, we conclude that any asiphonic CRN satisfies Assumption A.1(b). This definition coincides with the one of *exhaustive CRNs* introduced in [19].

Combining the following result with Remark 1.11 yields the LDP of Theorem 1.6 for the ASE networks of Assumption A.2.

PROPOSITION 1.12 (Existence of a Lyapunov function). *If the network is ASE, the generator $\mathcal{L}_v$ of* (1.2) *satisfies Assumption* A.1(*a*) *for the* chemical Lyapunov (*continuous*) function

$$U(x) := d + 1 + \sum_{i=1}^{d} x_i (\log x_i - 1) : \mathbb{R}_+^d \to \mathbb{R}_{\geq 1}. \tag{1.11}$$

The connection between Lyapunov stability analysis and large deviations rate functions is an active area of research (see, e.g., [5]). Also, the problem of stability of mass action kinetics systems has been addressed in [2, 12, 18, 20] and sufficient conditions for the existence of a globally attracting steady state for the deterministic dynamics of such systems have been established in [1, 7, 12]. In particular, the existence of a global attractor for a certain class of CRNs is proven in [1, 18] using the chemical Lyapunov function of (1.11). These results have been extended in [18] where the same function is used for showing the existence of a compact attracting set for strongly endotactic CRNs. However, none of the references above deal directly with the generator $\mathcal{L}_v$, using the chemical Lyapunov function to establish exponential tail estimates for the finite-time distributions of such stochastic processes, as we do in Section 3 [where we prove Proposition 1.12 by verifying (1.6) for this function].

REMARK 1.13. Proposition 1.12 implies that it is sufficient to check a set of graphical conditions to guarantee the applicability of a LDP to the dynamics of CRNs. This is most advantageous for applications in, for example, biochemistry, where typically $d > 100$ and quantitative estimates like (1.6) would be prohibitive to check. Note furthermore that our conditions do *not* depend on the reaction rate constants k_r, which are often very difficult to determine.

1.3. *Quasi-potential and exit time asymptotic*. Following the Wentzell–Freidlin approach, we utilize our sample path LDP to define the corresponding quasi-potential (as in [14]), and provide asymptotic analysis over an infinite time horizon, for quantities of interest such as the exit time from some domain $\mathcal{D} \subset \mathbb{R}_+^d$, or the transition time between different attractors of (1.4) (as proposed by [15]). To do so, we first assume that the domain of interest $\mathcal{D}$ has the following mild regularity properties.

ASSUMPTION A.3. The compact $\mathcal{D} \subset \mathbb{R}_+^d$ is the closure of its interior, with boundary $\partial\mathcal{D}$ that is a piecewise twice continuously differentiable submanifold of $\mathbb{R}_+^d$. Furthermore, there exists a ball $\mathcal{B} \subset \mathcal{D}$ so that for all $x \in \mathcal{B}$ and $y \in \mathcal{D}$ the set $\mathcal{D}$ contains the line segment between x and y.

DEFINITION 1.14. The *quasi-potential* between any $x, y \in \mathbb{R}_+^d$ is

$$\mathcal{V}_{\mathcal{D}}(x, y) := \inf_{t \geq 0} \inf_{\{z(\cdot) \in \mathcal{D}, z(t) = y\}} \{I_{x,t}(z)\},$$

for $I_{x,t}(z)$ of Theorem 1.6. We say that x, y are $\mathcal{D}$-equivalent (denoted $x \sim_{\mathcal{D}} y$), if $\mathcal{V}_{\mathcal{D}}(x, y) = \mathcal{V}_{\mathcal{D}}(y, x) = 0$. We further define

$$\mathcal{V}_{\mathcal{D}}(A, B) := \inf_{x \in A, y \in B} \mathcal{V}_{\mathcal{D}}(x, y) \qquad \forall A, B \subseteq \mathcal{D},$$

and use $\mathcal{V}(\cdot, \cdot)$ for $\mathcal{V}_{\mathbb{R}_+^d}(\cdot, \cdot)$.

The equivalence $x \sim_{\mathcal{D}} y$ defines compact sets $K_i \subset \mathcal{D}$ where the process can move with probability $\exp(-o(v))$. Throughout, we make the following assumption about their structure.

ASSUMPTION A.4 ([14], Condition A, Section 6.2). There exist ℓ compact sets $K_i \subset \mathcal{D}$ such that:

(a) every ω-limit set of (1.4) lying entirely in $\mathcal{D}$ is fully contained within one K_i,
(b) for any $x \in K_i$ we have $x \sim_{\mathcal{D}} y$ if and only if $y \in K_i$,
(c) for all K_i, the set K_j minimizing $\mathcal{V}_{\mathcal{D}}(K_i, K_j)$ is unique.

We further assume that the conic hull $\mathrm{Co}\{c^r\}_{r \in \mathcal{R}}$ of vectors $\{c^r\}_{r \in \mathcal{R}}$ is $\mathbb{R}^d$.

Such K_i are called *stable* if $\mathcal{V}(K_i^\delta, (K_i^\delta)^c) > 0$ for $\delta > 0$ small enough (where B^δ denotes the δ-neighborhood of the set B in the $\|\cdot\|_1$-norm). The most probable transitions between $\{K_i^\delta\}$ for small $\delta > 0$ and $v \to \infty$ [i.e., those transitions that connect any such K_i^δ to the unique K_j^δ with $i \neq j$ minimizing $\mathcal{V}_{\mathcal{D}}(K_i^\delta, K_j^\delta)$] define a deterministic dynamic on the finite collection of stable compact sets. Such dynamic can be partitioned into disjoint *cycles*, with each cycle π consisting of a single transitive point ($\pi = \{i\}$) or a periodic orbit $\pi = \{i_1 \to i_2 \to \cdots \to i_j \to i_1\}$ (cf. [14], Section 6.6, for the precise definition). Thanks to Assumptions A.3 and A.4, adapting the machinery of [14] to our setup, we transfer in Section 4 the sample path LDP to the following result about the time it takes the CRN to exit $\mathcal{D}$ or a cycle π and the probability cost of relevant exit paths.

THEOREM 1.15 ([14], Theorems 5.1, 5.3, 6.2, Section 6). *Consider a CRN satisfying Assumption A.1 and the process $t \mapsto X_t^v$ starting at $x_0^v \to x \in \mathcal{D}^o$. Let $\tau_{\mathcal{D}}$ denote its exit time from a set $\mathcal{D}$ that satisfies Assumptions A.3 and A.4 and let τ_π its first hitting time of $\bigcup_{j \notin \pi} K_j^\delta$ for a cycle $\pi \subseteq \{1, \dots, \ell\}$ and sufficiently small $\delta > 0$. Then, with nonrandom $M_{\mathcal{D}}(x)$, $W_{\mathcal{D}}$ and $W_{\mathcal{D}}(x, y)$ as in* [14], *Section* 6, *we have that for any x in a compact $F \subset \mathcal{D}^o$ and $y \in \partial\mathcal{D}$*

$$\lim_{\delta\to 0}\lim_{v\to\infty}\frac{1}{v}\log\mathbb{P}_{x_0^v}\big[\|X_{\tau_{\mathcal{D}}}^v - y\|_1 < \delta\big] = W_{\mathcal{D}} - W_{\mathcal{D}}(x, y), \tag{1.12}$$

$$\lim_{v\to\infty}\frac{1}{v}\log\mathbb{E}_{x_0^v}[\tau_{\mathcal{D}}] = W_{\mathcal{D}} - M_{\mathcal{D}}(x). \tag{1.13}$$

Furthermore, with $C(\pi) < \infty$ as in [14], *Section* 6.6, *any $\gamma > 0$ and uniformly in $x \in \bigcup_{i\in\pi}(K_i)^{\delta/2}$,*

$$\lim_{v\to\infty}\mathbb{P}_{x_0^v}\big[\big|v^{-1}\log\tau_\pi - C(\pi)\big| \le \gamma\big] = 1. \tag{1.14}$$

REMARK 1.16. Note that models in cell biology [4] usually have significantly larger dimension d than many other applications of Wentzell–Freidlin theory.

2. Proof of Theorem 1.6. We start by showing that Assumption A.1(a) yields exponentially negligible exit probability from the compact level sets of the function $U(\cdot)$.

LEMMA 2.1. *Let $\{X_t^v\}$ be a Markov jump process with generator* (1.2) *and initial condition $x_0^v \in (v^{-1}\mathbb{N}_0)^d$. Under Assumption* A.1(a), *there is, for every α, β, γ, a finite $\varrho_{\alpha,\beta,\gamma}$, so that*

$$\limsup_{v\to\infty}\frac{1}{v}\log\Big(\sup_{\|x_0^v\|_1\le\gamma}\mathbb{P}_{x_0^v}\Big[\sup_{t\in[0,e^{\beta v}]}\|X_t^v\|_1 > \varrho_{\alpha,\beta,\gamma}\Big]\Big) \le -\alpha. \tag{2.1}$$

PROOF. For each ℓ, there is a $\varrho = \varrho(\ell)$ so that $\{x : U(x) \le \ell\}$ is a subset of the ball

$$\widetilde{K}_\varrho := \big\{x \in \mathbb{R}_+^d : \|x\|_1 \le \varrho\big\}. \tag{2.2}$$

Considering the v-dependent stopping times

$$\sigma_\varrho := \inf\big\{t > 0 : X_t^v \notin \widetilde{K}_\varrho\big\}, \tag{2.3}$$

and the stopped processes $\widehat{X}_t^{v,\varrho} := X_{\sigma_\varrho \wedge t}^v$, it follows by Markov's inequality that for any T,

$$\mathbb{P}_{x_0^v}\Big[\sup_{t\in[0,T]} \|X_t^v\|_1 > \varrho\Big] = \mathbb{P}_{x_0^v}\big[\|\widehat{X}_T^{v,\varrho}\|_1 > \varrho\big]$$
$$\leq \mathbb{P}_{x_0^v}\big[U\big(\widehat{X}_T^{v,\varrho}\big) > \ell\big] \leq \ell^{-v}\mathbb{E}_{x_0^v}\big[U^v\big(\widehat{X}_T^{v,\varrho}\big)\big],$$

from which we get (2.1) once we show that

$$\sup_{\varrho\geq\gamma} \limsup_{v\to\infty} \frac{1}{v} \log \sup_{\|x_0^v\|_1\leq\gamma, T\leq e^{\beta v}} \mathbb{E}_{x_0^v}\big[U^v\big(\widehat{X}_T^{v,\varrho}\big)\big] < \infty. \tag{2.4}$$

To this end, as $U(\cdot)$ is continuous, $\sup_{\|x\|_1\leq\gamma}\{U(x)\} \leq e^{\varkappa}$ for some $\varkappa = \varkappa(\gamma) < \infty$. Further, when $\|X_0^v\|_1 \leq \varrho$, the Markov process $\widehat{X}_t^{v,\varrho}$ has the generator $\mathcal{L}_v$ of (1.2) restricted to $\widetilde{K}_\varrho$ and is confined for any $v \geq 1$ to a compact $(\widetilde{K}_\varrho)^{\bar{c}}$ with $\bar{c} := \sup_r \|c^r\|_1 < \infty$. Thus, combining Dynkin's formula [10], Section 5.1, with Assumption A.1(a) we find that for some $v_\varrho \in [1, \infty)$, all $v > v_\varrho$ and $\|x_0^v\|_1 \leq \gamma \leq \varrho$,

$$\mathbb{E}_{x_0^v}\big[U^v\big(\widehat{X}_T^{v,\varrho}\big)\big] \leq U^v\big(x_0^v\big) + \mathbb{E}_{x_0^v}\bigg[\int_0^{\sigma_\varrho\wedge T} \big(\mathcal{L}_v U^v\big)\big(X_s^v\big)\,\mathrm{d}s\bigg] \leq e^{\varkappa v} + T e^{bv}. \tag{2.5}$$

Considering for $T \leq e^{\beta v}$, the limit as $v \to \infty$ of v^{-1} times the logarithm of (2.5) leads to (2.4) and thereby concludes the proof. □

REMARK 2.2. Lemma 2.1 can be alternatively proved by defining a supermartingale from the condition (1.6) on our generator, and applying [13], Theorem 4.20, to it.

The Markov jump process X_{tT}^v corresponds to the generator of (1.2), now with reaction constants Tk_r for which Assumption A.1 continues to hold. This changes $\lambda(\cdot)$ of (1.5) to $T\lambda(\cdot)$, hence transforms $I_{x_0,T}(z(t))$ into $I_{x_0,1}(z(tT))$ [since $L(T\lambda, y) = TL(\lambda, T^{-1}y)$]. Thus, w.l.o.g., we take hereafter $T = 1$ and proceed to establish the exponential tightness of an exponentially equivalent process $\widetilde{X}_t^v$.

LEMMA 2.3. *Under Assumption* A.1(a), *the* $C_{0,1}(\mathbb{R}_+^d)$*-valued processes* $\widetilde{X}_t^v$ *obtained by linearly interpolating the jump points of* $t \mapsto X_t^v$, *form an exponentially tight family in the uniform topology, which for uniformly bounded* $\|x_0^v\|_1$ *is further exponentially equivalent to* $\{X_t^v\}$ *in the uniform topology on* $D_{0,1}(\mathbb{R}_+^d)$.

PROOF. For any consecutive jumps of X_t^v occurring at (random) times $t_1 < t_2$, we set

$$\widetilde{X}_t^v := X_{t_1}^v + \frac{t - t_1}{t_2 - t_1}\big(X_{t_2}^v - X_{t_1}^v\big).$$

Hence, $\|X_t^v - \widetilde{X}_t^v\|_1 \le v^{-1}\bar{c}$ for finite $\bar{c} := \sup_r \|c^r\|_1$, all $t \ge 0$, and v, yielding the exponential equivalence of $\{\widetilde{X}_t^v\}$ and $\{X_t^v\}$ (in the uniform topology). As for the exponential tightness of $\{\widetilde{X}_t^v\}$ in $C_{0,1}(\mathbb{R}_+^d)$, note that for any $t > s$,

$$\|\widetilde{X}_t^v - \widetilde{X}_s^v\|_1 \le v^{-1}\bar{c}N_{[s,t]}(X^v),$$

where $N_{[s,t]}(X^v)$ counts the number of jumps by $X_\cdot^v$ in the time interval $[s,t]$. Further, as $\Lambda_r^{(v)}(x) \le \lambda_r(x)$ for all $x \in \mathbb{R}_+^d$, we have for σ_ϱ of (2.3) and any $v \ge 1$ the monotone coupling $N_{[s,t]}(X^v) \le M_{[s,t]}^\varrho$ on $[0, \sigma_\varrho]$, where M^ϱ is a Poisson process of intensity $v\bar{\Lambda}^\varrho$ and

$$\bar{\Lambda}^\varrho := \sup_{x \in (\widetilde{K}_\varrho)^{\bar{c}}} \left\{\sum_{r\in\mathcal{R}} \lambda_r(x)\right\}.$$

In view of the Arzelà–Ascoli theorem and Lemma 2.1, it thus suffices for the stated exponential tightness of $\{\widetilde{X}_t^v\}$ to show that

$$\lim_{\delta\to 0} \limsup_{v\to\infty} v^{-1} \log \mathbb{P}\Big[\sup_{0\le s\le t\le (s+\delta)\wedge 1} \{M_{[s,t]}^\varrho\} \ge v\varepsilon\Big] = -\infty, \tag{2.6}$$
$$\forall \varrho < \infty, \varepsilon > 0.$$

To this end, by tail estimates for the Poisson$(2\delta v\bar{\Lambda}^\varrho)$ law, for any $\varepsilon > 0$ and $\varrho < \infty$,

$$\lim_{\delta\to 0} \limsup_{v\to\infty} v^{-1} \log \mathbb{P}[M_{[0,2\delta]}^\varrho \ge v\varepsilon] = -\infty.$$

Further, if $|t - s| \le \delta$ and $n = [1/\delta]$, then

$$M_{[s,t]}^\varrho \le \max_{i=0,\dots,n-1} \{M_{[i\delta,(i+2)\delta]}^\varrho\} =: \bar{M}_\delta^\varrho.$$

Hence, applying the union bound for the maximum $\bar{M}_\delta^\varrho$ of n identically distributed Poisson$(2\delta v\bar{\Lambda}^\varrho)$ variables yields (2.6), and thereby concludes the proof. □

Let $\mathcal{M}_1(\mathcal{S}_\star)$ denote the probability simplex over $\mathcal{S}_\star = \{\star\} \cup \mathcal{S}$ and $c_\star^r := \langle \mathbb{1}, c^r\rangle = \|c_{\text{out}}^r\|_1 - \|c_{\text{in}}^r\|_1$ the number of molecules gained (or lost, if negative) in each chemical reaction. For $\varrho > 0$ and $\{\lambda_r(\cdot)\}$ of (1.5) such that (1.4) admits a solution $\zeta : [0,1] \mapsto \mathbb{R}_+^d$ (i.e., no blowup on $[0,1]$), we consider $\mu(t)$ satisfying the ODE

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = \varrho^{-1} \sum_{r\in\mathcal{R}} \lambda_r(\varrho\mu|_{\mathcal{S}})(-c_\star^r, c^r), \qquad \mu(0) \in \mathcal{M}_1(\mathcal{S}_\star), \tag{2.7}$$

establishing a strictly positive lower bound on $\{\mu(t)|_{\mathcal{S}}\}$ that holds uniformly over $\|\mu(0)|_{\mathcal{S}}\|_1 \le \gamma/\varrho < 1$ with arbitrary, fixed γ and all ϱ large enough. This quantity is a rescaled projection on $\mathcal{M}_1(\mathcal{S}_\star)$ of the ODE (1.4) with initial condition $\|\zeta(0)\|_1 \le \gamma$ provided $\sup_t \|\zeta(t)\|_1 \le \varrho$. In other words, adding a "vacuum" species $\{\star\}$, we map $\zeta(t)$ onto $\mu(t)$, describing the dynamics a system conserving the total number of molecules. Note that $\mu(t)$ can be seen as the empirical measure of an IPS in the limit of infinite number of particles.

LEMMA 2.4. *Let Assumption A.1 hold and assume that* (1.4) *has a solution for* $t \in [0, 1]$. *Then, for any* $\gamma > 0$ *and for some* $\varrho_0(\gamma)$, *if* $\varrho \geq \varrho_0(\gamma)$ *and* $\varrho\mu(0)|_{\mathcal{S}} \leq \gamma$, *the solution* $\mu(t)$ *of* (2.7) *satisfies* $\mu(t) \in \mathcal{M}_1(\mathcal{S}_\star)$ *for* $t \in [0, 1]$.

Further, there exist $D \in \mathbb{N}$ *and* $b = b(\varrho) > 0$ *such that for any such* $\mu(t)$ *we have*

$$\mu_{s_i}(t) \geq bt^D \qquad \forall t \in [0, 1], i = 1, \ldots, d. \tag{2.8}$$

PROOF. Starting at $\langle \mathbb{1}, \mu(0)\rangle = 1$, it follows from the definition of $c^r_\star$ that $\langle \mathbb{1}, \mu(t)\rangle = 1$ for all $t \geq 0$, with the bijection

$$\zeta(t) = \varrho\mu(t)|_{\mathcal{S}} =: \Psi\bigl(\mu(t)\bigr), \qquad \mu_\star(t) := 1 - \varrho^{-1}\bigl\|\zeta(t)\bigr\|_1, \tag{2.9}$$

between $\mu(\cdot)$ of (2.7) and the assumed finite solution $\zeta(\cdot)$ of (1.4) with $\|\zeta(0)\|_1 \leq \gamma$. In particular, $\zeta(0) = \Psi(\mu(0)) \in \widetilde{K}_\varrho$ of (2.2) yields $\zeta(\cdot) \in \mathbb{R}^d_+$ and the condition $\varrho\mu(0)|_{\mathcal{S}} \leq \gamma$ translates into $\|\zeta(0)\|_1 \leq \gamma$. Our claim that $\mu(t) \in \mathcal{M}_1(\mathcal{S}_\star)$ for $t \in [0, 1]$ is thus just

$$\varrho_0(\gamma) := \sup_{\|\zeta(0)\|_1 \leq \gamma} \sup_{t \in [0,1]} \bigl\|\zeta(t)\bigr\|_1 < \infty,$$

which holds for $\varrho_0(\gamma) \leq 1 + \varrho_{1,0,\gamma}$ of (2.1) (indeed, simply contrast the FLLN of Remark 1.3 with the exponential decay in v of probabilities from Lemma 2.1).

Next, for any $\varrho > 0$ we multiply each reaction constant k_r by $\varrho^{\|c^r_{\text{in}}\|_1 - 1}$ and w.l.o.g. set hereafter $\varrho = 1$. Identifying $s_j = j$, split the RHS of (2.7) at coordinate i to a sum over reactions in $\mathcal{R}^i_+ := \{r \in \mathcal{R} : c^r_i > 0\}$ and over those in $\mathcal{R}^i_- := \{r \in \mathcal{R} : c^r_i < 0\}$. The contribution from $\mathcal{R}^i_+$ is a polynomial $P_i(\cdot)$ in $\{\mu_1, \ldots, \mu_d\}$ of positive coefficients (namely, $k_r c^r_i$, $r \in \mathcal{R}^i_+$). Further, $c^r_i < 0$ requires $(c^r_{\text{in}})_i \geq 1$ so the contribution of $\mathcal{R}^i_-$ is of the form $\mu_i Q_i(\mu)$ for another polynomial $Q_i(\cdot)$ with positive coefficients. Let $e(t) := \mu(t)|_{\mathcal{S}} - y(t)$, for the solution $y(t)$ of the modified ODEs

$$\frac{\mathrm{d}y_i}{\mathrm{d}t} = P_i\bigl(y(t)\bigr) - Cy_i(t), \qquad i = 1, \ldots, d, y(0) = \mu(0)|_{\mathcal{S}}, \tag{2.10}$$

where

$$C := 1 + \max_{i \leq d} \sup\bigl\{Q_i(\mu) : \mu \in \mathcal{M}_1(\mathcal{S}_\star)\bigr\} < \infty.$$

Each $P_i(\cdot)$ is increasing w.r.t. the natural partial order on $\mathbb{R}^d_+$, hence

$$\frac{\mathrm{d}e_i}{\mathrm{d}t} + Ce_i = P_i(y + e) - P_i(y) + \mu_i\bigl(C - Q_i(\mu)\bigr) \geq 0,$$

as long as $e(t)$ and $y(t)$ are both in $\mathbb{R}^d_+$, with a strict inequality as soon as $\mu_i(t) > 0$. Hence, starting at $e(0) = 0$ and $y(0) \in \mathbb{R}^d_+$, we establish (2.8) by showing that the same inequality holds if one substitutes the solution $y(\cdot)$ of (2.10) to $\mu(\cdot)$, uniformly over all $y(0) \in \mathbb{R}_+^d$. We achieve this goal by utilizing Assumption A.1(b) in at most d steps, to get that for some $D_k \in \mathbb{N}$ and $b_k > 0$,

$$y_i(t) \geq b_k t^{D_k}, \qquad \forall t \in [0,1], y(0) \in \mathbb{R}_+^d, i \in \mathcal{S}_k \uparrow \{1,\dots,d\}. \tag{2.11}$$

Specifically, starting at $\mathcal{S}_0 = \varnothing$ let $\mathcal{S}_k = \mathcal{S}_{k-1} \cup \partial\mathcal{S}_k$ for

$$\partial\mathcal{S}_k := \{ j \notin \mathcal{S}_{k-1} : \exists r \in \mathcal{R}, (c_{\text{out}}^r)_j > 0 \text{ and } \forall l \notin \mathcal{S}_{k-1}, (c_{\text{in}}^r)_l = 0 \}.$$

In particular, $\partial\mathcal{S}_1$ consists of all product species in reactions with $c_{\text{in}}^r = 0$ and from Assumption A.1(b) we know that $\partial\mathcal{S}_1$ is nonempty (see Remark 1.5). Such a reaction with $c_{\text{in}}^r = 0$ and an output $i \in \partial\mathcal{S}_1$ contributes to $P_i(\cdot)$ a positive constant term $k_{r,i} := k_r c_i^r$. For $y \in \mathbb{R}_+^d$ any other reaction may only increase $P_i(y)$, hence

$$\varkappa_1 := \inf_{i \in \partial\mathcal{S}_1} \inf_{y \in \mathbb{R}_+^d} \{P_i(y)\} > 0.$$

Bounding the solution of (2.10) from below taking $\varkappa_1$ instead of $P_i(y(t))$, and considering the worst case $y_i(0) = 0$, we deduce that for $k = 1$, $D_1 = 1$ and any $i \in \partial\mathcal{S}_k$,

$$y_i(t) \geq \varkappa_k \int_0^t e^{-C(t-s)} s^{D_k - 1} \, ds \geq b_k t^{D_k}, \tag{2.12}$$

for some $b_k = \varkappa_k g(C, D_k) > 0$ and $t \in [0,1]$. Increasing to $k = 2$, observe that if $\mathcal{S}_{k-1} \neq \mathcal{S}$ then by Assumption A.1(b) there must be a reaction r that has at least one product not from $\mathcal{S}_{k-1}$ while all of its substrates are from $\mathcal{S}_{k-1}$. In that case, the nonempty set $\partial\mathcal{S}_k$ consists of the products of such reactions that are not in $\mathcal{S}_{k-1}$ and for any $i \in \partial\mathcal{S}_k$ a reaction $r = r_i \in \mathcal{R}$ of this type contributes to $P_i(y(t))$ a positive term of the form

$$k_{r,i} \prod_{l \in \mathcal{S}_{k-1}} y_l(t)^{(c_{\text{in}}^r)_l} \geq k_{r,i} (b_{k-1} t^{D_{k-1}})^{\ell_i},$$

for $\ell_i := \|c_{\text{in}}^{r_i}\|_1$, where we relied on already having the bound (2.11) for $l \in \mathcal{S}_{k-1}$. Setting

$$D_k := 1 + D_{k-1} \max_{i \in \partial\mathcal{S}_k} \{\ell_i\}, \qquad \varkappa_k := \min_{i \in \partial\mathcal{S}_k} \{k_{r,i} b_{k-1}^{\ell_i}\},$$

recall that other reactions may only increase $P_i(y(t))$, hence for $i \in \partial\mathcal{S}_k$ and $t \in [0,1]$,

$$P_i(y(t)) \geq \varkappa_k t^{D_k - 1}.$$

Exactly as we have done for $k = 1$ and $D_1 = 1$, inserting such a lower bound into (2.10) and considering the worst case solution $[y_i(0) = 0]$, results with (2.12). Further lowering b_k to have the same bound extend also to all $i \in \mathcal{S}_{k-1}$ and proceeding if necessary to $k = 3$ and beyond exhausts finally all of $\mathcal{S}$ after at most d steps. □

PROOF OF THEOREM 1.6. Recall the Skorokhod J_1-topology on $D_{0,1}(\mathbb{R}^d_+)$ which is metrizable by the coarsening of the sup-norm

$$d_{J_1}(z_1, z_2) := \inf_{\tau}\Big\{\|\tau\|_\star + \sup_{s\in[0,1]} \big\|z_1(s) - z_2\big(\tau(s)\big)\big\|_1\Big\}, \tag{2.13}$$

where $\|\tau\|_\star := \sup_{s\neq t} \log\{|\tau(s) - \tau(t)|/|s - t|\}$ for strictly increasing $s \mapsto \tau(s)$ with $\tau(0) = 0$, $\tau(1) = 1$. By Lemma 2.3 and the inverse contraction principle of [8], Corollary 4.2.6, it suffices to establish the *weak* LDP for $\{\widetilde{X}^v_t\}$ in the metric space $(D_{0,1}(\mathbb{R}^d_+), d_{J_1})$ (in this standard reduction we also rely upon [8], Lemma 1.2.18, to upgrade from weak LDP to full LDP before employing the inverse contraction, and on [8], Theorem 4.2.13, to transfer the LDP in the uniform topology from $\{\widetilde{X}^v_t\}$ to $\{X^v_t\}$). Next, consider the Markov jump process $X^{v,\varrho}_t$ of generator (1.2) and volume-normalized jump rates

$$\Lambda^{v,\varrho}_r(x) := \Lambda^{(v)}_r(x)\mathbb{I}\big(\|x\|_1 \le \varrho - v^{-1}c^r_\star\big), \tag{2.14}$$

where $\mathbb{I}(A)$ is the indicator function over a set A. Taking $\bar{c}_\bullet := \sup_r \|c^r_{\text{in}}\|_1 \vee \|c^r_{\text{out}}\|_1$ and

$$\sup_{v\ge1}\{\|x^v_0\|_1\} + \bar{c}_\bullet \le \varrho, \tag{2.15}$$

assures that $\{X^{v,\varrho}_t, v \ge 1\}$ is confined to $\widetilde{K}_\varrho$ of (2.2) and in view of Lemma 2.1,

$$\lim_{\varrho\to\infty} \limsup_{v\to\infty} v^{-1} \log \mathbb{P}_{x^v_0}\big[X^{v,\varrho}_t \not\equiv X^v_t\big] = -\infty, \tag{2.16}$$

where $X^{v,\varrho}_t \not\equiv X^v_t$ represents the event where the paths of X^v_t and of $X^{v,\varrho}_t$ (coupled to X^v_t until the rates of the two processes differ) do not coincide on $t \in [0, 1]$. By taking τ as the identity map in (2.13), we further have for all v that

$$d_{J_1}\big(\widetilde{X}^v, X^v\big) \le \sup_{t\in[0,1]} \big\|\widetilde{X}^v_t - X^v_t\big\|_1 \le v^{-1}\bar{c}_\bullet$$

and consequently the required J_1-weak LDP for $\{\widetilde{X}^v_t\}$ follows from the *local* LDP for $\{X^v_t\}$ with respect to the d_{J_1}-metric balls (see [8], Theorem 4.1.11). In view of (2.16), the latter local LDP follows from having for any $z \in D_{0,1}(\mathbb{R}^d_+)$ and all ϱ large enough [which may depend on $z(\cdot)$],

$$\inf_{\delta>0} \limsup_{v\to\infty} \frac{1}{v} \log \mathbb{P}_{x^v_0}\big[d_{J_1}\big(X^{v,\varrho}, z\big) < \delta\big] \le -I_{x_0,1}(z), \tag{2.17}$$

$$\inf_{\delta>0} \liminf_{v\to\infty} \frac{1}{v} \log \mathbb{P}_{x^v_0}\big[d_{J_1}\big(X^{v,\varrho}, z\big) < \delta\big] \ge -I_{x_0,1}(z). \tag{2.18}$$

In establishing these bounds, we tackle the diminishing rates $\lambda_r(\cdot)$ at $\partial\mathbb{R}^d_+$ by employing a LDP from [9] for the empirical measure sample-path $t \mapsto \mu^n_t$ of n mean-field interacting particles. Specifically, fixing $z \in D_{0,1}(\mathbb{R}^d_+)$ let $\gamma := 1 + \sup_{t\in[0,1]} \|z(t)\|_1$. Since for any v and ϱ,

$$d_{J_1}\big(X^{v,\varrho}, z\big) < 1 \quad\Longrightarrow\quad \big\{X^{v,\varrho}_t : t \in [0, 1]\big\} \subseteq \widetilde{K}_\gamma, \tag{2.19}$$

the choice of jump rates $\Lambda^{v,\varrho}(\cdot)$ outside $\widetilde{K}_\gamma$ *is irrelevant* for the bounds (2.17) and (2.18). Choosing an integer ϱ with $\varrho \geq 2\varrho_0(\gamma) \geq 2\gamma$ which is further large enough for (2.15) to hold, the process $X^{v,\varrho}_\cdot$ is confined to $\widetilde{K}_\varrho$ of (2.2) so has at most $n = v\varrho$ molecules (to simplify notation, take w.l.o.g. $v \in \mathbb{N}$). We thus consider the evolution of n indistinguishable particles, each labeled by a type from $\mathcal{S}_\star$, where $n\mu^n_t(\star)$ counts the $\star$-particles that compensate the $c^r_\star$ molecules gained/lost at each reaction. Starting at $v(x^v_0)_i$ particles of type $s_i \in \mathcal{S}$ and $n - v\|x^v_0\|_1$ of $\star$-type, our goal is to have for $\Psi(\cdot)$ of (2.9) the continuous bijection

$$X^{v,\varrho}_t = \Psi(\mu^n_t), \qquad \mu^n_t(\star) = 1 - \varrho^{-1}\|X^{v,\varrho}_t\|_1. \tag{2.20}$$

To this end, a chemical reaction $r \in \mathcal{R}$ is mapped to the simultaneous change of $\ell_r := \|c^r_{\text{in}}\|_1 \vee \|c^r_{\text{out}}\|_1 \leq \bar{c}_\bullet$ particle types, where given μ^n, any ordered ℓ_r-tuple $\mathbf{i} \in \mathcal{S}^{\ell_r}_\star$ that has type-count configuration $((c^r_\star)_+, c^r_{\text{in}})$ independently changes into an ordered ℓ_r-tuple $\mathbf{j} \in \mathcal{S}^{\ell_r}_\star$ that has type-count configuration $((c^r_\star)_-, c^r_{\text{out}})$, at the rate

$$\Gamma^{(r),n}_{\mathbf{ij}}(\mu^n) = \frac{k_r \ell_r! v^{1-\|c^r_{\text{in}}\|_1}}{M_r \binom{n\mu^n(\star)}{(c^r_\star)_+}(c^r_\star)_+!}, \tag{2.21}$$

where $M_r = \ell_r!^2/[(|c^r_\star|)! \prod_{i=1}^d (c^r_{\text{in}})_i!(c^r_{\text{out}})_i!]$ is the number of pairs $(\mathbf{i}, \mathbf{j})$ matching the specified type-count configurations [and to accommodate all possible CRNs we permit $\mathbf{i}_l = \mathbf{j}_l$ for some l, unlike [9], equation (2.1)]. Indeed, for $\Lambda^{v,\varrho}_r$ of (2.14) and $\{\Gamma^{(r),n}_{\mathbf{ij}}\}$ of (2.21), the generator of μ^n in [9], equation (2.7), has total jump rate $v\Lambda^{v,\varrho}_r(\Psi(\cdot))$ in each direction $(-c^r_\star, c^r)$, $r \in \mathcal{R}$, thereby yielding the bijection property (2.20). From (2.21), it is also easy to check that for any $\mu^n \to \mu$,

$$n^{\ell_r-1}\Gamma^{(r),n}_{\mathbf{ij}}(\mu^n) \to \tilde{k}_r \mu_\star^{-(c^r_\star)_+} =: \Gamma^{(r)}_{\mathbf{ij}}(\mu),$$

where $\tilde{k}_r > 0$ is independent of μ. Such $\{\Gamma^{(r)}_{\mathbf{ij}}(\mu)\}$ satisfy the uniformity condition of [9], Assumption 3.1. On $\mathcal{M}_+(\mathcal{S}_\star) := \{\mu \in \mathcal{M}_1(\mathcal{S}_\star) : \mu_\star \geq 1/2\}$ they also have the Lipschitz continuity of [9], Assumption 2.2, and taking into account the factor v/n between volume normalizations, we have on $\mathcal{M}_+(\mathcal{S}_\star)$ the Lipschitz continuous asymptotic normalized reaction rates $\varrho^{-1}\lambda_r(\Psi(\mu))$ for μ^n that satisfy [9], Property 2.3. As shown in [9], Section 6, having [9], Property 2.3, throughout $\mathcal{M}_1(\mathcal{S}_\star)$ yields the LDP upper bound for $\{\mu^n_t\}$ in the J_1-topology of $D_{0,1}(\mathcal{M}_1(\mathcal{S}_\star))$, at rate n. Here, $\mu^n_0 \to \Psi^{-1}(x_0)$ and the asymptotic reaction rates for μ^n depend only on $\Psi(\mu)$. Consequently, the rate function controlling the LDP upper bound for $\{\mu^n(t)\}$ is

$$J(\mu) = \varrho^{-1} I_{x_0,1}(\Psi(\mu)),$$

and upon compensating for the factor v/n between the two rates, such an LDP upper bound for $\{\mu^n(t)\}$ readily yields (2.17). Our problem fails to satisfy the Lipschitz continuity of [9], Property 2.3, when $\mu_\star = 0$. However, $\varrho \geq 2\gamma$ guarantees that $\mu_\star \geq 1/2$ on $\Psi^{-1}(\widetilde{K}_\gamma)$, which in view of (2.19) is all that matters for (2.17). As explained at the start of the proof, upon combining (2.17) with the exponential tightness of Lemma 2.3, we get the stated LDP upper bound of (1.9) (for $T = 1$). In particular, due to exponential tightness the LHS of (1.9) is zero for some compact Γ. The same applies then for the infimum of $I_{x_0,1}(\cdot)$ over this compact set, and hence $I_{x_0,1}(\zeta) = 0$ for some $\zeta \in AC_{0,1}(\mathbb{R}_+^d)$ with $\zeta(0) = x_0$. Recall Remark 1.7 that such $\zeta(\cdot)$ must satisfy the ODE (1.4) for $t \in [0, 1]$. We note in passing that the same argument applies for any finite T (as explained just prior to Lemma 2.3), yielding the existence of global solutions for this ODE and further, the FLLN of Remark 1.3 then holds when $v_k/(\log k) \uparrow \infty$, by combining the stated LDP upper bound and the Borel–Cantelli lemma.

Next, note that we have (2.18) as a consequence of the LDP lower bound of [9] holding for $\{\mu_t^n\}$, equation (8.1). As mentioned in [9], Remark 8.6, such LDP applies when having in addition to [9], Property 2.3 and Assumption 3.1, also the η-ergodicity of [9], Assumption 3.3, and that the solution of the ODE (2.7) satisfies [9], Property 4.13. The latter amounts to having the lower bound of (2.8) also for $\mu_\star(t)$. Having $\varrho \geq 2\varrho_0(\gamma)$, from Lemma 2.4 this holds whenever starting at $\mu(0)|_{\mathcal{S}} \leq \gamma/\varrho$ which is precisely $\Psi^{-1}(\widetilde{K}_\gamma)$ [and thus all that is relevant for (2.18)]. The η-ergodicity of [9], Assumption 3.3, amounts here to being able to reach a particle population that exhibits all $d + 1$ elements of $\mathcal{S}_\star$ upon starting at $n \gg 1$ particles from a fixed, single type from $\mathcal{S}_\star$ and Assumption A.1(b) guarantees this when starting at only $\star$-particles. We thus have also [9], Assumption 3.3, except at the face $\mu_\star = 0$ on the boundary of $\mathcal{M}_1(\mathcal{S}_\star)$. While the behavior at that face plays a role for some events, thanks to (2.19) it is irrelevant here. □

3. The stability of ASE networks. The proof of Proposition 1.12 is long and technically challenging, so we first sketch in Section 3.1 the proof of (1.6) for x away from $\partial\mathbb{R}_+^d$ to familiarize the reader with the techniques used in the subsequent sections, where we carry out this proof in full detail.

3.1. *Toric rays and outline of the stability proof.* Following the geometrical analysis of [18], we first define *toric rays*, using throughout for $w \in \mathbb{R}^n$, $z \in (\mathbb{R}_+^n)^o$ and $\theta \in \mathbb{R}_+$ the operators

$$\log(z) := (\log z_1, \dots, \log z_n) \in \mathbb{R}^n,$$

$$z^w := (z_1^{w_1}, \dots, z_n^{w_n}) \in (\mathbb{R}_+^n)^o,$$

$$\theta^w := (\theta^{w_1}, \dots, \theta^{w_n}) \in (\mathbb{R}_+^n)^o.$$

DEFINITION 3.1. To each w in the unit sphere S^{n-1}, we associate the *w-toric ray*

$$T^w = \bigcup_{\theta \geq 1} \theta^w \subset \mathbb{R}_+^n.$$

We also introduce the toric ray parameters

$$\theta(z) := \exp(\|\log(z)\|_2), \qquad w(z) := \frac{1}{\log\theta(z)}\log(z), \tag{3.1}$$
$$(\theta, w) : (\mathbb{R}_+^n)^o \to \mathbb{R}_{>1} \times S^{n-1}, \qquad z = \theta(z)^{w(z)}.$$

REMARK 3.2. To see why $U(\cdot)$ of (1.11) is most suitable for mass action systems, note that along a w-toric ray

$$\nabla U(\theta^w) = \log(\theta^w) = (\log\theta)w, \tag{3.2}$$

while the derivative of the ODE (1.4) at a point on such a ray is

$$\frac{d\zeta}{dt}\bigg|_{\zeta=\theta^w} = \sum_{r\in\mathcal{R}} \lambda_r(x)c^r\bigg|_{\zeta=\theta^w} = \sum_{r\in\mathcal{R}} k_r(\theta^w)^{c_{\text{in}}^r}c^r = \sum_{r\in\mathcal{R}} k_r\theta^{\langle w, c_{\text{in}}^r\rangle}c^r. \tag{3.3}$$

Thus, at $x = \theta^w$ the time derivative of $U(\zeta(t))$ for the solution $\zeta(t)$ of (1.4) is

$$\frac{d}{dt}U(\zeta(t))\bigg|_{\zeta=\theta^w} = \left\langle \nabla U(x), \frac{d\zeta}{dt}\right\rangle\bigg|_{\zeta=\theta^w} = (\log\theta)\sum_{r\in\mathcal{R}} k_r\langle w, c^r\rangle\theta^{\langle w, c_{\text{in}}^r\rangle}. \tag{3.4}$$

For fixed w and $\theta \gg 1$, the sum on the RHS of (3.4) is dominated by reactions $r \in \mathcal{R}_w$ (maximizing $\langle w, c_{\text{in}}^r\rangle$). Thus, in strongly endotactic CRNs, where at least one such reaction contributes negatively to this sum by having $\langle w, c^r\rangle < 0$, and no other reaction r in $\mathcal{R}_w$ contributes positively to it, the LHS of (3.4) will also be negative for all large enough θ. As shown in [18], if this applies uniformly over $w \in S^{d-1}$ then for some compact K we have $\frac{d}{dt}U(\zeta(t)) < 0$ whenever $\zeta(t) \notin K$, so (1.4) has an absorbing compact set. Indeed, suppose to the contrary, that for some diverging sequence $x(j) \in \mathbb{R}_+^d$

$$\frac{d}{dt}U(\zeta(t))\bigg|_{\zeta=x(j)} \geq 0 \qquad \forall j \in \mathbb{N}. \tag{3.5}$$

By compactness of S^{d-1}, upon passing to a suitable subsequence, the corresponding toric ray parameters $x(j) = \theta(j)^{w(j)}$ form a *toric jet* of frame $\bar{w} = \{\bar{w}^{(k)} : k \leq \ell\}$ (see Definition 3.11 and [18], Lemma 6.7), where $w(j) \to \bar{w}^{(1)}$ and $\theta(j) \to \infty$. By compactness of $[1, \infty]$, there exists a further subsequence along which $x(j)^{\bar{w}^{(k)}}$ converge for each k (possibly to ∞), implying the convergence of the functions $\widehat{\varphi}_r(x) := k_r\langle w, c^r\rangle\theta^{\langle w, c_{\text{in}}^r\rangle}$. For strongly endotactic CRNs, one can show [18] that along such a toric jet, for any $r \in \mathcal{R}$ there exists $r' \in \mathcal{R}$ whose contribution $\widehat{\varphi}_{r'}(x(j))$ to the RHS of (3.4) is such that $\lim_j \widehat{\varphi}_{r'}(x(j)) < 0$ and $-\widehat{\varphi}_{r'}(x(j))/(\widehat{\varphi}_r(x(j)))_+ \to \infty$ (where $0^{-1} := \infty$), contradicting (3.5).

REMARK 3.3. Note that in the components $s_i \in \mathcal{S}$ where w_i is negative, the value of θ^{w_i} decreases as $\theta \to \infty$. Such w are therefore used to parametrize through (3.1) points in $(v^{-1}\mathbb{N}_0)^d$ that are at a distance < 1 from the boundary $\{x_i = 0\}$.

Proposition 1.12 amounts to having for some finite b, for any $x \in (v^{-1}\mathbb{N}_0)^d$ and for $v > v'(\|x\|_1)$,

$$\sum_{r\in\mathcal{R}} \Lambda_r^{(v)}(x)\big[U^v\big(x+v^{-1}c^r\big) - U^v(x)\big] \le e^{bv}. \tag{3.6}$$

Recall that $\Lambda_r^{(v)}(x) \le \lambda_r(x)$ which is uniformly bounded on compacts, as is $U(x)$. Hence, there exists a finite $b = b(\varrho)$ such that (3.6) holds for any $v \ge 1$ whenever $\|x\|_1 \le \varrho$. Letting

$$\mathcal{A}^v_{\varrho,\varrho'} := \{x \in (v^{-1}\mathbb{N}_0)^d : \varrho < \|x\|_1 \le \varrho'\},$$
$$L_r^{(v)}(x) := U(x)\big(Q_r^{(v)}(x) - 1\big), \qquad Q_r^{(v)} := U^v\big(x+v^{-1}c^r\big)/U^v(x),$$

we thus establish Proposition 1.12 upon showing that for some $\varrho < \infty$ any $\varrho' \ge \varrho$, $x \in \mathcal{A}^v_{\varrho,\varrho'}$ and $v > v'(\varrho')$, we have

$$a^{(v)}(x) := \sum_{r\in\mathcal{R}} k_r \varphi_r^{(v)}(x) \le 0, \qquad \varphi_r^{(v)}(x) := k_r^{-1}\Lambda_r^{(v)}(x) L_r^{(v)}(x), \tag{3.7}$$

where, by (1.3) one considers in $a^{(v)}(x)$ only r such that $vx \ge c^r_{\text{in}}$ (thus $x + v^{-1}c^r \in \mathbb{R}^d_+$). Subject to having the v-independent approximation for all $x \in (v^{-1}\mathbb{N})^d$,

$$Q_r^{(v)}(x) = \exp\left[\frac{h_r(x) + \mathcal{O}_{\|x\|}(1)}{U(x)}\right] \qquad \text{with } h_r(x) := \langle \nabla U(x), c^r\rangle, \tag{3.8}$$

we can prove (3.7), at least for a strictly positive x, by contradiction. Specifically, one can show that it suffices to rule out having $a^{(v(j))}(x(j)) > 0$ along a *rapidly diverging* volume-jet $(v(j), x(j))$. That is, along some diverging toric jet $x(j) = \theta(j)^{w(j)} \in (v(j)^{-1}\mathbb{N})^d$, with $\theta(j) \to \infty$ and frame $\bar{w}$, such that $v(j) \to \infty$ arbitrarily fast [i.e., allowing for an arbitrary $v'(\varrho)$ in Definition 3.13]. Similar to Remark 3.2, we arrive at a contradiction by showing that for some v' any such v'-divergent volume-jet (v, x) and $r \in \mathcal{R}$, there must exist $r' \in \mathcal{R}_{\bar{w}^{(1)}}$ such that eventually $\varphi_{r'}^{(v)}(x) < 0$ and $-\varphi_{r'}^{(v)}(x)/(\varphi_r^{(v)}(x))_+ \to \infty$. To this end, we first show in Lemma 3.14 that for $v'(\varrho) = e^{\varrho}$ and some constants $\delta_{r'} > 0$, along any v'-divergent volume-jet (v, x) framed by $\bar{w}$, eventually

$$\Lambda_{r'}^{(v)}(x) \ge \delta_{r'}\lambda_{r'}(x) \qquad \forall r' \in \mathcal{R}_{\bar{w}^{(1)}}, \tag{3.9}$$

which as $\Lambda_r^{(v)}(x) \le \lambda_r(x)$, implies that for $C < \infty$, any $r \in \mathcal{R}$ and $r' \in \mathcal{R}_{\bar{w}^{(1)}}$, eventually

$$C\left|\frac{\varphi_{r'}^{(v)}(x)}{\varphi_r^{(v)}(x)}\right| \ge \frac{k_{r'}^{-1}\lambda_{r'}(x)}{k_r^{-1}\lambda_r(x)}\left|\frac{L_{r'}^{(v)}(x)}{L_r^{(v)}(x)}\right| =: P_{r,r'}^{(v)}(x). \tag{3.10}$$

Referring to the first part in the RHS of (3.10) as a *monomial term* [since $k_{r'}^{-1}\lambda_{r'}(x)/k_r^{-1}\lambda_r(x) = \theta^{\langle w, c^{r'}_{\text{in}} - c^r_{\text{in}}\rangle}$], and to the second part (in the absolute value sign) as *Lyapunov term*, we then show that for any $r \in \mathcal{R}$, if eventually $L_r^{(v)}(x) > 0$ then by [18], Proposition 6.24, there exists $r' \in \mathcal{R}_{\bar{w}^{(1)}}$ with $h_{r'}(x) \to -\infty$ such that along the divergent volume-jet,

$$\lim_{j\to\infty} P_{r,r'}^{(v(j))}\big(x(j)\big) = \infty. \tag{3.11}$$

Indeed, relying on (3.8) we establish (3.11) by proceeding according to whether $\varkappa_r := \lim_j |h_r(x)|$ is finite or infinite. Specifically, we have the following cases:

(a) Lyapunov domination, where $\varkappa_r$ is finite and with $U(x) \to \infty$ we have that $L_r^{(v)}(x)$ remains uniformly bounded. The existence of $r'(r) \in \mathcal{R}_{\bar{w}^{(1)}}$ with $L_{r'}^{(v)}(x) \to -\infty$ resulting from [18], Proposition 6.24 (see Lemma 3.17), combined with [18], Lemma 6.22, to bound the monomial term away from zero, concludes the proof.

(b) Monomial domination, where $\varkappa_r = \infty$ so $Q_r^{(v)}(x) = e^{h_r(x)/U(x)}(1 + o(1))$ by (3.8). This implies, by [18], Propositions 6.20 and 6.24, the existence of $r' \in \mathcal{R}_{\bar{w}^{(1)}}$ such that $|L_{r'}^{(v)}(x)/L_r^{(v)}(x)| = \mathcal{O}(\theta^{-\langle w,c^r\rangle/U(x)})$, whose exponent goes to zero as $j \to \infty$. On the other hand, for such r' by [18], Lemma 6.22, the exponent of θ in the monomial term of (3.10) is (eventually) strictly positive and bounded away from zero along the toric jet, thereby establishing (3.11).

In order to establish (3.7) also on $\partial\mathbb{R}_+^d$, we need to extend the preceding program to deal with boundary effects such as the divergence of $\nabla U(x)$. This is done by separately considering each face of $\mathbb{R}_+^d$. In particular, Section 3.2 establishes (3.8) in a more general form, substituting $\nabla U(\cdot)$ with the *v-dependent* $\nabla_r^{(v)} U(\cdot)$ of (3.12). Section 3.3 adapts the definitions of toric jet and strongly endotactic CRNs from [18] as needed for $\partial\mathbb{R}_+^d$. This and the corresponding results from [18], Section 6, are then used in Section 3.4 to show the divergence of $P_{r,r'}^{(v)}(x)$, first for Lyapunov domination (in Lemma 3.21), and then for monomial domination (in Lemma 3.22). Finally, Section 3.5 follows the preceding outline in combining everything to a proof of Proposition 1.12.

3.2. *Approximation lemmas*. The image of $\mathbb{R}^d$ under the exponential map is $(\mathbb{R}_+^d)^o$, so we will establish (3.7) separately on the various faces of $\partial\mathbb{R}_+^d$ by considering the relevant CRNs $(\mathcal{S}, \mathcal{C}, \mathcal{R}(\mathcal{P}))$ where, for any nonempty $\mathcal{P} \subseteq \mathcal{S}$,

$$\mathcal{R}(\mathcal{P}) := \{r \in \mathcal{R} : \operatorname{supp}\{c_{\text{in}}^r\} \subseteq \mathcal{P}\}$$

denotes the reactions with substrates only from $\mathcal{P}$. To this end, identify such $\mathcal{P} = (s_{i_1}, \dots, s_{i_{d_\mathcal{P}}})$ of cardinality $|\mathcal{P}| = d_\mathcal{P} \geq 1$ with the corresponding indices $(i_1, \dots, i_{d_\mathcal{P}})$, denoting by $\mathbb{S}^d(\mathcal{P})$ the restriction of a space $\mathbb{S}^d$ (be it $\mathbb{R}^d$, $\mathbb{R}_+^d$, $\mathbb{N}_0^d$ or $\mathbb{N}^d$), to these coordinates (i.e., having zero values outside $\mathcal{P}$). Aiming at the approximation (3.8) for $x \in (v^{-1}\mathbb{N})^d(\mathcal{P})$ and $r \in \mathcal{R}(\mathcal{P})$, we modify $\nabla U(x)$ to

$$(\nabla_r^{(v)} U(x))_i := \begin{cases} \log x_i, & i \in \mathcal{P}, \\ \log(v^{-1} c_i^r), & i \in \operatorname{supp}\{c_{\text{out}}^r\} \cap \mathcal{P}^c, \\ 0 & \text{otherwise}. \end{cases} \tag{3.12}$$

We write $\varepsilon_v(x)$ for functions that are uniformly bounded in x by some $\bar{\varepsilon}_v \to 0$ as $v \to \infty$ and $\varepsilon(x)$ for any globally bounded function of x.

LEMMA 3.4. *Setting $g_p(x) := \|\log(x)\|_p$ for $p = 1, 2$, we have that*

$$\frac{2g_1(x)}{vU(x)} \le \frac{d + g_2(x)^2}{vU(x)} \le \varepsilon_v(x) \qquad \forall x \in (v^{-1}\mathbb{N})^d. \tag{3.13}$$

PROOF. The first inequality in (3.13) directly results from the fact that $x^2 + 1 \ge 2|x|$ for all $x \in \mathbb{R}$. Next, since $g_2(x) \le \sqrt{d} \sup_i\{|\log x_i|\}$ and $U(x) \ge 1$, by (1.11) it suffices to show that

$$\frac{|\log y|^2}{v[y(\log y - 1) + 2]} \le \varepsilon_v(y) \qquad \forall y \ge v^{-1}.$$

For $y \in [v^{-1}, v]$, the LHS is at most $(\log v)^2/v \to 0$ as $v \to \infty$, whereas for $y \ge v \ge e^2$ the LHS is bounded above by $2\log y/(vy) \le 2\log v/v^2 \to 0$ as $v \to \infty$. $\square$

REMARK 3.5. Hereafter, for any $r \in \mathcal{R}(\mathcal{P})$, we consider w.l.o.g. only *r-relevant* x, namely those for which $vx + c^r \in \mathbb{N}_0^d$, for otherwise the corresponding term disappears in (3.7) (see Remark 1.2).

LEMMA 3.6. *There is a finite v_* such that for any $\mathcal{P} \subseteq \mathcal{S}$, all $r \in \mathcal{R}(\mathcal{P})$ and all r-relevant $x \in (v^{-1}\mathbb{N})^d(\mathcal{P})$, one has for $v \ge v_*$:*

$$Q_r^{(v)}(x) = \exp\left[\frac{h_r^{(v)}(x) + \varepsilon(x)}{U(x)}\right],$$

where $h_r^{(v)}(x) := \langle \nabla_r^{(v)} U(x), c^r \rangle$.

PROOF. Since the number of possible $\mathcal{P}$ and r is finite, it suffices to prove the claim for fixed $\mathcal{P}$ and r. We have in terms of $f := v[U(x + v^{-1}c^r) - U(x)]/U(x)$ that

$$Q_r^{(v)}(x) = \left(1 + \frac{f}{v}\right)^v = \exp[f - vR(f/v)]$$

where the nonnegative $R(y) := y - \log(1 + y)$ satisfies

$$R(y) \le 2y^2 \qquad \forall y \ge -1/2. \tag{3.14}$$

Now, for any $r \in \mathcal{R}(\mathcal{P})$ and $x \in (v^{-1}\mathbb{N})^d(\mathcal{P})$ with $vx + c^r \in \mathbb{N}_0^d$ we have that

$$fU(x) - h_r^{(v)}(x) = \sum_{i \in \mathcal{P}} \psi(vx_i; c_i^r) - \langle c^r, \mathbb{1} \rangle = \varepsilon(x) \tag{3.15}$$

is uniformly bounded since $\psi(b; c) := (b + c)\log(1 + c/b)$ decreases in $b \geq \max(1, -c)$. Hence, from (3.15), (3.12) and Lemma 3.4,

$$\frac{1}{2} f^2 \leq \left(\frac{h_r^{(v)}(x)}{U(x)}\right)^2 + \left(\frac{\varepsilon(x)}{U(x)}\right)^2 = \frac{v\varepsilon_v(x)}{U(x)}.$$

Since $U(x) \geq 1$, we see that $(f/v)^2 \leq 2\varepsilon_v(x)/v \leq 1/4$ for some v_* finite, any $v \geq v_*$ and all x, in which case by (3.14) we have that $vR(f/v) \leq 2f^2/v \leq 4\frac{\varepsilon_v(x)}{U(x)}$, as claimed. □

3.3. *Strongly endotactic subnetworks and divergent volume-jets.* Throughout, for nonempty $\mathcal{P} \subseteq \mathcal{S}$ and $w \in \mathbb{R}^d$ we denote by $\pi_\mathcal{P} : \mathbb{R}^d \to \mathbb{R}^{d_\mathcal{P}}$ the projection onto the coordinates with indices in $\mathcal{P}$. Proceeding to adapt for $(\mathcal{S}, \mathcal{C}, \mathcal{R}(\mathcal{P}))$ key definitions from CRN theory, such as *strongly endotactic* (see [18]), for all $w \in \mathbb{R}^d$ with nonzero projection $w_\mathcal{P} := \pi_\mathcal{P} w$, let $\mathcal{R}(\mathcal{P})_w$ denote the reactions having c_{in}^r in the w-maximal subset of $\mathcal{C}_{\text{in}}(\mathcal{P}) = \{c_{\text{in}}^r : r \in \mathcal{R}(\mathcal{P})\}$. Clearly, $\mathcal{R}(\mathcal{P})_w$ depends only on $w_\mathcal{P}$ which w.l.o.g. is in the $(d_\mathcal{P} - 1)$-dimensional unit sphere $S^\mathcal{P}$ and w.l.o.g. we further identify $\mathcal{C}_{\text{in}}(\mathcal{P})$ with $\pi_\mathcal{P}\mathcal{C}_{\text{in}}(\mathcal{P})$.

DEFINITION 3.7. Fixing $w_\mathcal{P} \in S^\mathcal{P}$, a reaction $r \in \mathcal{R}(\mathcal{P})$ with $\text{supp}\{c_{\text{out}}^r\} \subseteq \mathcal{P}$ is called *w-dissipative*, *w-null* or *w-explosive* according to $\langle w, c^r \rangle = \langle w_\mathcal{P}, \pi_\mathcal{P} c^r \rangle$ being negative, zero or positive, respectively. Any $r \in \mathcal{R}(\mathcal{P})$ having some product species not within $\mathcal{P}$ is considered *w-dissipative* (regardless of w). Similarly, $r \in \mathcal{R}(\mathcal{P})$ is $\{w\}$-explosive, $\{w\}$-null or $\{w\}$-dissipative, if the relevant property holds for all but finitely many elements of $\{w\} \subset S^\mathcal{P}$.

REMARK 3.8. For $\mathcal{P} = \mathcal{S}$ our Definition 3.7 of w-dissipative and w-explosive reactions, coincides with [18], Definition 6.15, of w-sustaining and w-draining reactions, respectively. The nomenclature was changed to stress the behavior of reactions for $\|x\|_1 \gg 1$ which is of interest here: dissipative (explosive) reactions contribute to the decrease (increase) of the Lyapunov function along trajectories far away from the origin.

We next extend Definition 1.8 of strongly $\mathcal{S}$-endotactic CRN to $\mathcal{P} \subset \mathcal{S}$. Such an extension is needed in light of Remark 3.2, and made relevant by Lemma 3.10.

DEFINITION 3.9. For any $w \in \mathbb{R}^d$ with nonzero projection onto $\mathcal{P}$ (or $w_\mathcal{P} \in S^\mathcal{P}$), the CRN $(\mathcal{S}, \mathcal{C}, \mathcal{R}(\mathcal{P}))$ is called *w-strongly $\mathcal{P}$-endotactic* if the set $\mathcal{R}(\mathcal{P})_w$ contains at least one w-dissipative reaction, and no w-explosive reactions. Such a CRN is called *strongly $\mathcal{P}$-endotactic* if it is w-strongly $\mathcal{P}$-endotactic for all w as above.

LEMMA 3.10. *Any strongly endotactic* $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ *is strongly* $\mathcal{P}$*-endotactic for all* $\mathcal{P} \subset \mathcal{S}$ *if* $\mathcal{R}(\mathcal{P}) \neq \varnothing$.

PROOF. Fixing $\mathcal{P} \subset \mathcal{S}$ with $\mathcal{R}(\mathcal{P}) \neq \varnothing$, suppose that for a nonzero w there is a w-explosive $r \in \mathcal{R}(\mathcal{P})_w$. Since the nonnegative $\sum_{i \notin \mathcal{P}} (c_{\text{in}}^{r'})_i$ is zero iff $r' \in \mathcal{R}(\mathcal{P})$, setting $w_i' = w_i$ for $i \in \mathcal{P}$ and $w_i' = -\gamma$ for $i \notin \mathcal{P}$ we have that $\mathcal{R}_{w'} = \mathcal{R}(\mathcal{P})_w$ for γ large enough. Further, $\text{supp}\{c^r\} \subseteq \mathcal{P}$ hence $\langle c^r, w' \rangle = \langle c^r, w \rangle > 0$, so having $r \in \mathcal{R}_{w'}$ contradicts Definition 1.8. For the same reason, if every reaction in $\mathcal{R}(\mathcal{P})_w$ is w-null, then for large γ the same applies for every reaction in $\mathcal{R}_{w'}$, in contradiction with Definition 1.8. □

To show that (3.7) holds *whenever* $v > v'(\|x\|_1)$ and $vx \in \mathbb{N}^d(\mathcal{P})$ with $\|x\|_1 \geq \varrho$, requires an approximation framework for sequences $x(j) = \theta(j)^{w(j)}$ satisfying $\theta(j) \to \infty$ and $w(j) \to \bar{w}^{(1)}$ in $S^{\mathcal{P}}$. To this end, we follow [18], Section 6, in coding the latter convergence by a suitable $d_\star$-dimensional *frame* [23]: an orthonormal system (ONS) $\bar{w} := \{\bar{w}^{(1)}, \dots, \bar{w}^{(d_\star)}\} \subset S^{\mathcal{P}}$ such that

$$\lim_{j \to \infty} \frac{\beta^{(k+1)}}{\beta^{(k)}} = 0 \qquad \forall k < d_\star, \qquad \beta^{(k)} = \beta^{(k)}(j) := \langle w(j), \bar{w}^{(k)} \rangle. \tag{3.16}$$

For generic $\{w(j)\}$, one needs a full $d_{\mathcal{P}}$-dimensional basis of $S^{\mathcal{P}}$, but degeneracy allows for $d_\star < d_{\mathcal{P}}$ [e.g., $\bar{w}^{(1)}$ alone suffices when all $w(j)$ lie on a single toric ray]. Further, the *order* within $\bar{w}$ is adapted to the sequence, so that the angle between $w(j)$ and $\bar{w}^{(k)}$ decays faster with each increase of k. Through the following definition, by slight abuse of notation we suppress the index j for elements of the sequence $\{x(j)\}$ and other related quantities to increase the readability of forthcoming formulas.

DEFINITION 3.11 ([18], Definitions 6.2, 6.18). (a) A *unit jet* on a frame $\bar{w}$ is a sequence $\{w\} = \{w(j)\}$ of unit vectors in the conic hull $\text{Co}(\bar{w})$ satisfying (3.16).

(b) A *toric jet* $\{x\}$ is a sequence $\theta(j)^{w(j)} \in \mathbb{R}^d_{>0}(\mathcal{P})$ for a unit jet $\{w\}$ and $\theta(j) \to \infty$.

(c) A unit jet $\{w\}$ and the corresponding toric jets are *adapted* to $(\mathcal{S}, \mathcal{C}, \mathcal{R}(\mathcal{P}))$ if the classification of each $r \in \mathcal{R}(\mathcal{P})$ according to Definition 3.7 is conserved by all $w(j)$ with $j \in \mathbb{N}$ and for all $k = 1, \dots, d_\star$, $\lim_j \theta^{\beta^{(k)}}$ exists and takes values in $[1, \infty]$.

REMARK 3.12. When the unit jet $\{w\}$ is adapted to $(\mathcal{S}, \mathcal{C}, \mathcal{R}(\mathcal{P}))$ and clear from the context, in view of point (c) we call a reaction $r \in \mathcal{R}(\mathcal{P})$ dissipative or explosive, per Definition 3.7, without explicitly indicating the choice of $w(j)$.

Having assigned any $r \in \mathcal{R}(\mathcal{P})$ with $\text{supp}\{c_{\text{out}}^r\} \nsubseteq \mathcal{P}$ as dissipative reactions, it is necessary for the strategy presented in Section 3.1 to ensure that their contribution to $a^{(v)}(x)$ is negative along $\{(v, x)\}$ in case $r \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$. Since for such a reaction $\lim_v \langle \nabla_r^{(v)} U(x), c^r \rangle = -\infty$, we obtain in Lemma 3.16 the desired behavior of $L_r^{(v)}(x)$ by choosing, for every x to have $v > v'(\|x\|_1)$ for a function $v'(\cdot)$ increasing fast enough. Our next definition guarantees that this condition on v is met along $\{x\}$.

DEFINITION 3.13. Fixing $\mathcal{P} \subseteq \mathcal{S}$ and an increasing function $v'(\varrho)$, we call a sequence of $\{(v, x) : vx \in \mathbb{N}^d(\mathcal{P}), v > v'(\|x\|_1)\}$ a $(v', \mathcal{P})$-*divergent volume-jet* if $\{x\}$ is a toric jet for a unit jet $\{w\}$ framed by $\bar{w}$ that is adapted to $(\mathcal{S}, \mathcal{C}, \mathcal{R}(\mathcal{P}))$, such that $\lim_{j\to\infty} \|x\|_1 = \infty$.

As we show next, using this framework further yields the estimate (3.9) [which, as outlined in Section 3.1, is the first step in proving (3.7)].

LEMMA 3.14. *Setting* $v'(\varrho) = e^{\varrho}$, *there exists* $\delta > 0$ *such that for any frame* $\bar{w}$, $r \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ *and* $(v', \mathcal{P})$-*divergent volume-jet* (v, x) *framed by* $\bar{w}$, *eventually,*

$$\lambda_r(x) \geq \Lambda_r^{(v)}(x) \geq \delta \lambda_r(x). \tag{3.17}$$

PROOF. Letting $\xi(j) := j! j^{-j}$ for $j \in \mathbb{N}$ and $\xi(0) = 1$, we set

$$\delta_r := \prod_{i=1}^{d} \xi\big((c_{\text{in}}^r)_i\big) > 0.$$

As mentioned before, comparing (1.3) and (1.5) one gets the first inequality of (3.17) for any $x \in (v^{-1}\mathbb{N}_0)^d$, $v \geq 1$. Further, the ratio $\Lambda_r^{(v)}(x)/\lambda_r(x)$ is nondecreasing in each vx_i and equals δ_r when $vx = c_{\text{in}}^r$. Thus, setting $\delta = \min_r \delta_r$ it suffices to show that for $r \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ and a $(v', \mathcal{P})$-divergent volume-jet $\{(v, x)\}$ framed by $\bar{w}$, we eventually have $vx_i \geq (c_{\text{in}}^r)_i$. This trivially holds if $(c_{\text{in}}^r)_i = 0$, so the proof is complete upon showing that, along $\{(v, x)\}$,

$$i \in \operatorname{supp}\{c_{\text{in}}^r\} \quad \Longrightarrow \quad \lim_{j\to\infty} \{\log v + w_i \log \theta\} = \infty. \tag{3.18}$$

Since $(\log \|\pi_{\mathcal{P}} x\|_1)/(\log \theta) \to \max_i \{\bar{w}_i^{(1)}\} =: \psi$ and both $\|x\|_1$ and θ diverge, we have $\psi \geq 0$. Further, $w_i \to \bar{w}_i^{(1)}$ is finite and $\log v \geq \log v'(\|x\|_1) = \|x\|_1$ so (3.18) clearly holds whenever $\psi > 0$. In case $\psi = 0$, the vector $\bar{w}^{(1)} \in S^{\mathcal{P}}$ has nonpositive coordinates, so $\bar{w}_{i'}^{(1)} \leq -1/\sqrt{d}$ for some $i' \in \mathcal{P}$. Since $\lim_j w = \bar{w}^{(1)}$, it then follows that eventually $w_{i'} \leq -1/\sqrt{2d} =: -\zeta$. Since $i' \in \mathcal{P}$ and $vx \in \mathbb{N}^d(\mathcal{P})$, this implies that $v \geq \theta^{-w_{i'}} \geq \theta^{\zeta}$. Recall Remark 1.5 that some $r' \in \mathcal{R}(\mathcal{P})$ has $c_{\text{in}}^{r'} = 0$, hence $\langle \bar{w}^{(1)}, \pi_{\mathcal{P}} c_{\text{in}}^r \rangle \geq 0$ for any $r \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$. That is, when $i \in \operatorname{supp}\{c_{\text{in}}^r\}$ we have that $w_i \to \bar{w}_i^{(1)} = 0$ and as $\log v \geq \zeta \log \theta$, we recover (3.18) and with it, complete the proof. □

Finally, adapting [18], Definitions 6.8, 6.15, each possible frame within $S^{\mathcal{P}}$, induces two key indices (classifications) for reactions $r \in \mathcal{R}(\mathcal{P})$.

DEFINITION 3.15. For $\mathcal{P} \subseteq \mathcal{S}$ and ordered ONS $\bar{w} \subset S^{\mathcal{P}}$: (a) Let $\text{super}_1 = \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ and define for $k \geq 2$ the nested subsets super_k of reactions having c^r_{in} in the $\bar{w}^{(k)}$-maximal subset of $\{\pi_{\mathcal{P}} c^r_{\text{in}} : r \in \text{super}_{k-1}\}$. (b) The *level* ℓ within $\bar{w}$ of $r \in \mathcal{R}(\mathcal{P})$ having $\text{supp}\{c^r_{\text{out}}\} \subseteq \mathcal{P}$ is $\ell := \inf\{k : \langle \bar{w}^{(k)}, \pi_{\mathcal{P}} c^r \rangle \neq 0\}$ (with $\ell = \infty$ when no such k exists), setting $\ell = 1$ if r has some product species outside $\mathcal{P}$.

3.4. *The dominance of dissipative reactions*. Turning to the proof of (3.11), we first bound (in the setting of Lemma 3.14), the contribution to the Lyapunov term when $r' \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ and $\text{supp}\{c^{r'}_{\text{out}}\} \nsubseteq \mathcal{P}$, allowing us thereafter to simultaneously treat such reactions and those in $\mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ with $\text{supp}\{c^{r'}_{\text{out}}\} \subseteq \mathcal{P}$, ultimately using their negative contribution to dominate any positive term in $a^{(v)}(x)$ from (3.7).

LEMMA 3.16. *For $v'(\varrho) = e^{\varrho}$ and any $(v', \mathcal{P})$-divergent volume-jet (v, x) framed by $\bar{w}$:*

(a) *If $r \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ and $\text{supp}\{c^r_{\text{out}}\} \nsubseteq \mathcal{P}$, then*

$$\limsup_{j \to \infty} \left\{ \frac{h^{(v)}_r(x)}{\log \theta} \right\} < 0. \tag{3.19}$$

(b) *If $r \in \mathcal{R}(\mathcal{P})$ has $\text{supp}\{c^r_{\text{out}}\} \subseteq \mathcal{P}$ then $\varkappa_r = \infty$ iff r has finite level ℓ and $\lim_j \theta^{\beta^{(\ell)}} = \infty$.*

PROOF. (a) Recall that $h^{(v)}_r(x) = \langle \nabla^{(v)}_r U(x), c^r \rangle$, so setting $\alpha_r := \langle \mathbb{I}_{\mathcal{P}^c}, c^r_{\text{out}} \rangle > 0$ and $\eta_r := \langle \mathbb{I}_{\mathcal{P}^c} \log c^r_{\text{out}}, c^r_{\text{out}} \rangle$ which is finite, we have from (3.12) that

$$\frac{h^{(v)}_r(x)}{\log \theta} = \frac{\eta_r}{\log \theta} + \langle w, \pi_{\mathcal{P}} c^r \rangle - \alpha_r \frac{\log v}{\log \theta}. \tag{3.20}$$

Because $\theta = \theta(j) \to \infty$, the first term on the RHS decays to zero and the second term converges to $\langle \bar{w}^{(1)}, \pi_{\mathcal{P}} c^r \rangle$. While proving (3.18), we have seen that if $\psi := \max_i \{\bar{w}^{(1)}_i\} > 0$, then $\log v \geq \|x\|_1$ (for the v'-divergent volume-jet), results with $(\log v)/(\log \theta) \to \infty$ and consequently (3.19) holds. In case $\psi = 0$, we have shown in that same proof that $(\log v)/(\log \theta) \geq \zeta > 0$ along the divergent volume-jet and further that $\langle \bar{w}^{(1)}, \pi_{\mathcal{P}} c^r_{\text{in}} \rangle = 0$ when $r \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$. Recall that $c^r = c^r_{\text{out}} - c^r_{\text{in}}$ with $c^r_{\text{out}} \in \mathbb{R}^d_+$ and $\psi = 0$ amounts to $-\bar{w}^{(1)} \in \mathbb{R}^{d_{\mathcal{P}}}_+$. Thus, in this setting $\langle \bar{w}^{(1)}, \pi_{\mathcal{P}} c^r \rangle \leq 0$, which by (3.20) recovers (3.19).

(b) If $r \in \mathcal{R}(\mathcal{P})$ has $\text{supp}\{c^r_{\text{out}}\} \subseteq \mathcal{P}$, then $h^{(v)}_r(\cdot) = h_r(\cdot)$ is independent of v and in (3.20) we have $\alpha_r = \eta_r = 0$. Further, recall Definition 3.11 that $\{w\} \subset \text{Co}(\bar{w})$, so if r has infinite level then $h_r(\cdot) = 0$, while if it has finite level ℓ within $\bar{w}$, then by (3.16), along the divergent volume-jet

$$\lim_{j\to\infty} \frac{h_r^{(v)}(x)}{\beta^{(\ell)}\log\theta} = \langle \bar{w}^{(\ell)}, \pi_{\mathcal{P}} c^r \rangle \neq 0, \tag{3.21}$$

from which the stated criterion for divergence of $|h_r^{(v)}(x)|$ follows. $\square$

Our next result shows that $L_{r'}^{(v)}(x) \to -\infty$ for any dissipative $r' \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ with $\varkappa_{r'} = \infty$ (see Section 3.1 for explanation about the Lyapunov domination).

LEMMA 3.17. *For $v'(\varrho) = e^{\varrho}$, if $r \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ with $\varkappa_r = \infty$ is dissipative for a $(v', \mathcal{P})$-divergent volume jet (v, x) framed by $\bar{w}$, then*

$$\lim_{j\to\infty} L_r^{(v)}(x) = -\infty.$$

PROOF. By Definition 3.13, the toric jet $\{x\}$ is adapted to $(\mathcal{S}, \mathcal{C}, \mathcal{R}(\mathcal{P}))$. Hence, if $\mathrm{supp}\{c_{\mathrm{out}}^r\} \subseteq \mathcal{P}$ and $r \in \mathcal{R}(\mathcal{P})$ is dissipative, then by Definition 3.7 and (3.2),

$$h_r^{(v)}(x) = h_r(x) = (\log\theta)\langle w, \pi_{\mathcal{P}} c^r \rangle < 0 \qquad \forall j.$$

Since $\varkappa_r = \infty$, it follows that $h_r^{(v)}(x) \to -\infty$ as $j \to \infty$, which by part (a) of Lemma 3.16 applies also when $\mathrm{supp}\{c_{\mathrm{out}}^r\} \nsubseteq \mathcal{P}$. Fixing $\gamma < \infty$, since $\varepsilon(x)$ of Lemma 3.6 is uniformly bounded, we thus have that for all j large enough,

$$L_r^{(v)}(x) = U(x)\big(Q_r^{(v)}(x) - 1\big) \leq U(x)\big[e^{-\frac{\gamma}{U(x)}} - 1\big] \leq -\gamma + \frac{\gamma^2}{2U(x)}$$

(as $e^{-y} \leq 1 - y + \frac{y^2}{2}$ for $y \in \mathbb{R}_+$). Recalling from Definition 3.13 that $\|x(j)\|_1 \to \infty$, and consequently $U(x(j)) \to \infty$, we complete the proof by taking $j \to \infty$ followed by $\gamma \to \infty$. $\square$

We plan to show that if $r \in \mathcal{R}(\mathcal{P})$ has $L_r^{(v)}(x) > 0$ along some $(v', \mathcal{P})$-divergent volume-jet $\{(v, x)\}$ for $v'(\varrho) > e^{\varrho}$, then (3.11) holds for a $\{w\}$-dissipative $r' \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$. To this end, we first introduce the CRN $\mathcal{C}_{\bar{w}^{(1)},\mathcal{P}}$ in which necessarily $\mathrm{supp}\{c_{\mathrm{out}}^r\} \subseteq \mathcal{P}$ [or else by Lemma 3.16(a) and Lemma 3.17 eventually $L_r^{(v)}(x) < 0$].

DEFINITION 3.18. For $\mathcal{P} \subseteq \mathcal{S}$ and $u \in S^{\mathcal{P}}$, let $(\mathcal{S}, \mathcal{C}_{u,\mathcal{P}}, \mathcal{R}(\mathcal{P}))$ denote the CRN obtained upon restricting c_{out}^r to $\mathbb{R}_+^d(\mathcal{P})$ for any $r \notin \mathcal{R}(\mathcal{P})_u$.

REMARK 3.19. Of course, $\mathcal{C}_{u,\mathcal{P}} = \mathcal{C}$ when $\mathcal{P} = \mathcal{S}$. More generally, this modification never affects $\{c_{\mathrm{in}}^r\}$, hence neither the rates $\Lambda_r^{(v)}(\cdot)$ nor the sets $\{\mathcal{R}(\mathcal{P})_w, w \in S^{\mathcal{P}}\}$, or super_k of Definition 3.15. Further, the CRN $(\mathcal{S}, \mathcal{C}_{u,\mathcal{P}}, \mathcal{R}(\mathcal{P}))$

remains u-strongly $\mathcal{P}$-endotactic (see Definition 3.9), and thus also $w(j)$-strongly $\mathcal{P}$-endotactic, for j large enough and any unit jet $\{w(j)\}$ whose frame starts at $\bar{w}^{(1)} = u$.

Comparing our Definition 3.11 and Definition 3.15 with the corresponding definitions of [18], Section 6, it is easy to verify that [18], Theorem 6.11, Lemmas 6.7, 6.10, 6.19, apply in our setting as does [18], Proposition 6.20.1 (for draining reactions), even for the modified CRN of Definition 3.18. We next adapt to the latter setting, those conclusions of [18], Lemma 6.22, Proposition 6.24, that we shall use in the sequel.

LEMMA 3.20. *Fix a strongly endotactic CRN* $(\mathcal{S}, \mathcal{C}, \mathcal{R})$. *Consider the corresponding CRN of Definition* 3.18 *and ordered ONS* $\bar{w}$ *of length* ℓ' *starting at* $\bar{w}^{(1)} = u$. *Then*:

(a) *If* $\operatorname{supp}\{c^r_{\text{out}}\} \subseteq \mathcal{P}$, $\langle \bar{w}^{(k)}, \pi_{\mathcal{P}} c^r \rangle = 0$ *for* $k < \ell'$ *and* $\langle \bar{w}^{(\ell')}, \pi_{\mathcal{P}} c^r \rangle > 0$, *then* $r \notin \text{super}_{\ell'}$.

(b) *Some* $r' \in \text{super}_{\ell'}$ *has* $\operatorname{supp}\{c^{r'}_{\text{out}}\} \nsubseteq \mathcal{P}$ *or* $k \mapsto \langle \bar{w}^{(k)}, \pi_{\mathcal{P}} c^{r'} \rangle$ *not identically zero*, *with a negative first nonzero term*.

PROOF. Since $k \mapsto \text{super}_k$ are nested sets, it suffices to rule out that respectively:

(a$'$) Some $r \in \text{super}_{\ell'}$ has $\operatorname{supp}\{c^r_{\text{out}}\} \subseteq \mathcal{P}$, $\langle \bar{w}^{(k)}, \pi_{\mathcal{P}} c^r \rangle = 0$ for $k < \ell'$ and $\langle \bar{w}^{(\ell')}, \pi_{\mathcal{P}} c^r \rangle > 0$.

(b$'$) Each $r \in \text{super}_{\ell'}$ has $\operatorname{supp}\{c^r_{\text{out}}\} \subseteq \mathcal{P}$ and $\langle \bar{w}^{(k)}, \pi_{\mathcal{P}} c^r \rangle = 0$ for all $k \le \ell'$.

Further, the modification of Definition 3.18 neither affects $\text{super}_{\ell'}$ nor the value of c^r for reactions in $\text{super}_{\ell'} \subseteq \mathcal{R}(\mathcal{P})_u$ (see Remark 3.19), so it suffices to rule out (a$'$) and (b$'$) for $(\mathcal{S}, \mathcal{C}, \mathcal{R}(\mathcal{P}))$ and the given ONS $\bar{w}$. To this end, consider a unit jet $\{w\}$ framed by $\bar{w}$, adapted to $(\mathcal{S}, \mathcal{C}, \mathcal{R}(\mathcal{P}))$ and having $\beta^{(\ell')}(j) > 0$ for all j. Recall [18], Theorem 6.11, that $\mathcal{R}(\mathcal{P})_{w(j)} = \text{super}_{\ell'}$ eventually in j. Thus, by (3.16), our assumptions (a$'$) respectively (b$'$) imply that for all large enough j, respectively:

(a†) There exists a $w(j)$-explosive $r \in \mathcal{R}(\mathcal{P})_{w(j)}$ of level ℓ'.

(b†) The collection $\mathcal{R}(\mathcal{P})_{w(j)}$ consists of only $w(j)$-null reactions.

To conclude, note that (a†) and (b†) contradict having a strongly $\mathcal{P}$-endotactic $(\mathcal{S}, \mathcal{C}, \mathcal{R}(\mathcal{P}))$. $\square$

Similar to [18], Proposition 6.26, we proceed via a pair of lemmas that establish (3.11) for $(\mathcal{S}, \mathcal{C}_{\bar{w}^{(1)},\mathcal{P}}, \mathcal{R}(\mathcal{P}))$ by bounding from below the asymptotic behavior of the Lyapunov and monomial terms, as in cases (a) and (b) at the end of Section 3.1, that correspond to $\varkappa_r < \infty$ and $\varkappa_r = \infty$, respectively.

LEMMA 3.21 (Lyapunov domination). *For* $v'(\varrho) = e^{\varrho}$ *and the ONS* $\bar{w}$ *for* $\mathcal{P} \subseteq \mathcal{S}$, *consider the CRN* $(\mathcal{S}, \mathcal{C}_{\bar{w}^{(1)},\mathcal{P}}, \mathcal{R}(\mathcal{P}))$ *and a* $(v', \mathcal{P})$*-divergent volume jet*

(v, x) for it, framed by $\bar{w}$. Then, for any $r \in \mathcal{R}(\mathcal{P})$ with $\mathrm{supp}\{c^r_{\mathrm{out}}\} \subseteq \mathcal{P}$ *and $\varkappa_r < \infty$, the domination* (3.11) *holds for some dissipative $r' \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$.*

PROOF. Let ℓ denote the level of $r \in \mathcal{R}(\mathcal{P})$ within the frame $\bar{w}$, if finite, whereas if the level of r is infinite, set $\ell = d_\star + 1$ and $\beta^{(\ell)} \equiv 0$. Since $\mathrm{supp}\{c^r_{\mathrm{out}}\} \subseteq \mathcal{P}$, we have from Lemma 3.16 (b) that $\lim_j \theta^{\beta^{(\ell)}} < \infty$. For any divergent volume-jet $\beta^{(1)} \to 1$, hence $\lim_j \theta^{\beta^{(1)}} = \infty$, $\ell \geq 2$ and in view of (3.16) there exists $1 \leq \ell' < \ell$ such that

$$\lim_{j\to\infty} \theta^{\beta^{(\ell')}} = \infty \quad \text{and} \quad \lim_{j\to\infty} \theta^{\beta^{(\ell'+1)}} < \infty. \tag{3.22}$$

For the subframe $\{\bar{w}^{(1)}, \dots, \bar{w}^{(\ell')}\}$, Lemma 3.20(b) yields $r' \in \mathrm{super}_{\ell'} \subseteq \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ of level $\ell_\star \leq \ell'$ such that either $\mathrm{supp}\{c^{r'}_{\mathrm{out}}\} \nsubseteq \mathcal{P}$ or $\langle \bar{w}^{(\ell_\star)}, \pi_\mathcal{P} c^{r'} \rangle < 0$. Since $\lim_j \beta^{(k)}/\beta^{(k+1)} = \infty$ for any $k \geq 1$, such r' must also be $\{w\}$-dissipative. Proceeding to establish (3.11), by Lemma 3.6 combined with $e^{|x|} - 1 \leq 2|x|$ for $|x| < 1$ and $h^{(v)}_r(x)/U(x) \to 0$, we have for j large enough $|L^{(v)}_r(x)| \leq 2(|\varepsilon(x)| + |h^{(v)}_r(x)|) \leq C^{-1}$, hence

$$P^{(v)}_{r,r'}(x) \geq C\theta^{\langle w, \pi_\mathcal{P}(c^{r'}_{\mathrm{in}} - c^r_{\mathrm{in}})\rangle} |L^{(v)}_{r'}(x)|.$$

As $r' \in \mathrm{super}_{\ell'}$ and $c^r_{\mathrm{in}} \in \mathcal{C}_{\mathrm{in}}(\mathcal{P})$ we have from [18], Lemma 6.10.2, that for any $k_\star \leq \ell' + 1$, $\delta > 0$ and all j large enough

$$\begin{aligned}\langle w(j), \pi_\mathcal{P}(c^{r'}_{\mathrm{in}} - c^r_{\mathrm{in}})\rangle &\geq \sum_{k=k_\star}^{d_\star} \beta^{(k)}(j) \langle \bar{w}^{(k)}, \pi_\mathcal{P}(c^{r'}_{\mathrm{in}} - c^r_{\mathrm{in}})\rangle \\ &\geq \beta^{(k_\star)}(j)[\langle \bar{w}^{(k_\star)}, \pi_\mathcal{P}(c^{r'}_{\mathrm{in}} - c^r_{\mathrm{in}})\rangle - \delta].\end{aligned} \tag{3.23}$$

Taking $k_\star = \ell' + 1$ [where if $\ell' = d_\star$ then $\langle \bar{w}^{(k)}, \pi_\mathcal{P}(c^{r'}_{\mathrm{in}} - c^r_{\mathrm{in}})\rangle \geq 0$ for all $k \leq d_\star$ hence the LHS of (3.23) is nonnegative], we deduce from (3.22) that $\theta^{\langle w, \pi_\mathcal{P}(c^{r'}_{\mathrm{in}} - c^r_{\mathrm{in}})\rangle}$ is uniformly (in j) bounded below. The proof is thus complete upon showing that $\varkappa_{r'} = \infty$, as then $|L^{(v)}_{r'}(x)| \to \infty$ by Lemma 3.17. Indeed, since $r' \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ from Lemma 3.16(a) we have that $\varkappa_{r'} = \infty$ if $\mathrm{supp}\{c^{r'}_{\mathrm{out}}\} \nsubseteq \mathcal{P}$, whereas if $\mathrm{supp}\{c^{r'}_{\mathrm{out}}\} \subseteq \mathcal{P}$ then r' of finite level $\ell_\star \leq \ell'$ has $\varkappa_{r'} = \infty$ in view of the LHS of (3.22) and Lemma 3.16(b). □

LEMMA 3.22 (Monomial domination). *For $v'(\varrho) = e^\varrho$ and ONS $\bar{w}$ for $\mathcal{P} \subseteq \mathcal{S}$, consider the CRN $(\mathcal{S}, \mathcal{C}_{\bar{w}^{(1)},\mathcal{P}}, \mathcal{R}(\mathcal{P}))$ and a $(v', \mathcal{P})$-divergent volume jet (v, x) for it, framed by $\bar{w}$. Then, for any $\{w\}$-explosive $r \in \mathcal{R}(\mathcal{P})$ with* $\mathrm{supp}\{c^r_{\mathrm{out}}\} \subseteq \mathcal{P}$ *and $\varkappa_r = \infty$, the domination* (3.11) *holds for some dissipative $r' \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$.*

PROOF. By Lemma 3.16(b), here r has finite level ℓ' within $\bar{w}$ for which the LHS of (3.22) holds. Further, with $\{w(j)\}$ adapted to $(\mathcal{S}, \mathcal{C}_{\bar{w}^{(1)},\mathcal{P}}, \mathcal{R}(\mathcal{P}))$ we deduce from [18], Proposition 6.20.1, that since $\langle w(j), \pi_{\mathcal{P}} c^r \rangle$ is positive for j large, $\langle \bar{w}^{(\ell')}, \pi_{\mathcal{P}} c^r \rangle$ must also be positive, hence Lemma 3.20(a) yields that $r \notin \mathrm{super}_{\ell'}$. Recall the proof of Lemma 3.21, that there exists $\{w\}$-dissipative $r' \in \mathrm{super}_{\ell'} \subseteq \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$. In particular, $\langle \bar{w}^{(\ell')}, \pi_{\mathcal{P}}(c_{\mathrm{in}}^{r'} - c_{\mathrm{in}}^{r}) \rangle$ is positive, so considering (3.23) for $k_\star = \ell'$ and small $\delta > 0$, for j large enough we bound the monomial term of (3.10) by

$$\theta^{\langle w, \pi_{\mathcal{P}}(c_{\mathrm{in}}^{r'} - c_{\mathrm{in}}^{r}) \rangle} \geq \left(\theta^{\beta^{(\ell')}}\right)^{\delta}. \tag{3.24}$$

Further, the $\{w\}$-dissipative $r' \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ has level $\ell_\star \leq \ell'$ and $\varkappa_{r'} = \infty$, hence

$$K_{r'} := \lim_{j \to \infty} \frac{h_{r'}^{(v)}(x)}{\beta^{(\ell')} \log \theta} \tag{3.25}$$

is strictly negative [see (3.19) for $\mathrm{supp}\{c_{\mathrm{out}}^{r'}\} \not\subseteq \mathcal{P}$ and (3.21) otherwise]. The $\{w\}$-explosive r has level ℓ' and $\varkappa_r = \infty$ hence by (3.21) it satisfies (3.25) for some $0 < K_r < \infty$. Recall (3.22) that $\beta^{(\ell')} \log \theta$ diverges along our jet $\{w\}$. Hence, by Lemma 3.6, for any $s > K_r$ and $\gamma \in (0, 1)$ such that $\gamma s < -K_{r'}$, the corresponding Lyapunov term is eventually bounded below by

$$\frac{1 - Q_{r'}^{(v)}(x)}{Q_r^{(v)}(x) - 1} \geq \frac{1 - (\theta^{-s\beta^{(\ell')}/U(x)})^{\gamma}}{\theta^{s\beta^{(\ell')}/U(x)} - 1} \geq \gamma \theta^{-s\beta^{(\ell')}/U(x)} \tag{3.26}$$

[where the second inequality follows from $1 - \xi^{\gamma} \geq \gamma(1 - \xi)$ which holds for any $\xi, \gamma \in (0, 1)$]. With $U(x) \to \infty$, the RHS of (3.24) dominates the RHS of (3.26) and the divergence of $P_{r,r'}^{(v)}(x)$ of (3.10) follows. □

3.5. *Proof of Proposition* 1.12. By (3.7), Proposition 1.12 will hold if we can find $\varrho < \infty$ such that for any $\varrho' < \infty$,

$$v'(\varrho') := \sup\Big\{v : \sup_{x \in \mathcal{A}_{\varrho,\varrho'}^{v}} \{a^{(v)}(x)\} > 0\Big\} < \infty.$$

Assume to the contrary, that there exist $\varrho'_j > \varrho_j \uparrow \infty$, $v(j,k) \to \infty$ as $k \to \infty$ and $x(j,k) \in \mathcal{A}_{\varrho_j,\varrho'_j}^{v(j,k)}$ such that $a^{v(j,k)}(x(j,k)) > 0$ for all $j, k \in \mathbb{N}^2$. Then, for any desired increasing $v'(\cdot)$, upon choosing $k = k_j$ large enough, we extract a sequence $\{(v(j), x(j))\}$ such that $v(j)x(j) \in \mathbb{N}_0^d$, $\|x(j)\|_1 \to \infty$ and

$$a^{v(j)}\big(x(j)\big) > 0, \qquad v(j) > v'\big(\|x(j)\|_1\big) \qquad \forall j \in \mathbb{N}. \tag{3.27}$$

Since $d < \infty$, there must be some $\mathcal{P} \subseteq \mathcal{S}$ such that $v(j)x(j) \in \mathbb{N}^d(\mathcal{P})$ along some infinite subsequence. Also, as $\|x(j)\|_1 \to \infty$, upon restriction to $(\mathcal{S}, \mathcal{C}, \mathcal{R}(\mathcal{P}))$ we have that $\theta(x(j)) \to \infty$ [see (3.1)], and our Definition 3.11 of unit jet and toric jet then coincide with those of [18]. Hence, by [18], Lemma 6.7, we extract a subsubsequence $(v(j), x(j))$ satisfying all of the above, for which in addition $\{x(j)\}$

is a toric jet for a unit jet $\{w(j)\}$ framed by some $\bar{w}$. Finally, in view of [18], Lemma 6.19, there exists a further sub-sub-subsequence $\{x(j)\}$ which is adapted to $(\mathcal{S}, \mathcal{C}, \mathcal{R}(\mathcal{P}))$ [note that $\text{supp}\{c^r_{\text{out}}\} \nsubseteq \mathcal{P}$ has nothing to do with the choice of $\{w(j)\}$]. In conclusion, we have a $(v', \mathcal{P})$-divergent volume-jet $\{(v, x)\}$ satisfying (3.27), where we are free to choose $v'(\varrho)$ and only $r \in \mathcal{R}(\mathcal{P})$ is to be considered in (3.7). Fixing $\{(v, x)\}$ and in particular its frame $\bar{w}$, we may and can move to the CRN $(\mathcal{S}, \mathcal{C}_{\bar{w}^{(1)},\mathcal{P}}, \mathcal{R}(\mathcal{P}))$ of Definition 3.18. Indeed, recall Remark 3.19 that this does not affect the rates $\Lambda^{(v)}_r(x)$, while for $v \geq \sup_r \|c^r_{\text{out}}\|_\infty$ and $x \in \mathbb{R}^d_+(\mathcal{P})$ it may only increase $L^{(v)}_r(x)$, by setting to zero some negative contributions $(v^{-1}c^r)_i[\log(v^{-1}c^r)_i - 1]$ to $U(x + v^{-1}c^r)$ from $i \in \text{supp}\{c^r_{\text{out}}\} \setminus \mathcal{P}$. As explained before, in the new CRN $\text{supp}\{c^r_{\text{out}}\} \nsubseteq \mathcal{P}$ requires $r \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ and further subsampling our divergent volume-jet to make it adapted to $(\mathcal{S}, \mathcal{C}_{\bar{w}^{(1)},\mathcal{P}}, \mathcal{R}(\mathcal{P}))$, we proceed as outlined in Section 3.1 to show that on the latter CRN, having (3.27) leads to a contradiction. Indeed, consider $r \in \mathcal{R}(\mathcal{P})$, whose contribution to (3.27) is eventually positive (for the modified reactions of $\mathcal{C}_{\bar{w}^{(1)},\mathcal{P}}$). That is, having $L^{(v)}_r(x) > 0$ for all j large. By Lemma 3.6, this requires $h^{(v)}_r(x) + \varepsilon(x) > 0$, which in view of Lemma 3.16(a) implies that $\text{supp}\{c^r_{\text{out}}\} \subseteq \mathcal{P}$. With $\{x\}$ adapted, this yields, as in the proof of Lemma 3.16(b), that $|h^{(v)}_r(x)| \to \varkappa_r$ when $j \to \infty$ [see (3.21)], and further that $\varkappa_r = \infty$ is possible only for a $\{w\}$-explosive reaction. For both $\varkappa_r < \infty$ and $\varkappa_r = \infty$, we now have (3.11) for some dissipative $r' \in \mathcal{R}(\mathcal{P})_{\bar{w}^{(1)}}$ (see Lemma 3.21 and Lemma 3.22, resp.). As (3.10) is a consequence of Lemma 3.14, it follows that $a^{(v)}(x) \leq 0$ along $\{(v, x)\}$, in contradiction with (3.27).

4. Proof of Theorem 1.15. Theorem 1.15 is proved in [14], Section 6, for a uniformly elliptic diffusion on a compact d-dimensional manifold, when the driving Brownian motion has been scaled by ε. Recall that such a diffusion satisfies an LDP with rate $v := \varepsilon^{-2}$ and its good rate function is zero iff $x'(t) = b(x(t))$ starting at $x(0) = x_0$. We have here the analogous LDP of Theorem 1.6, whose good rate function is zero iff $z(t)$ solves the ODE (1.4) (see Remark 1.7). Further, with our Assumptions A.4 and A.3 replacing [14], Condition A, Section 6.2, and [14], Section 6.5, respectively, we merely adapt the proof in [14], Section 6, where the stated results are established from [14], Lemmas 6.1.1–6.1.9. Specifically, for (1.12) and (1.13) which concern only the dynamics of $t \mapsto X^v_t$ within the compact $\mathcal{D}$, it suffices that we prove the weaker version Lemma 4.1 of [14], Lemma 6.1.1, within $\mathcal{D}$, and the modification Lemma 4.2 of [14], Lemma 6.1.4, while tackling the degeneracy of $\{X^v_t\}$ on $\partial\mathbb{R}^d_+$. Indeed, Lemma 4.1 and Lemma 4.2 suffice for establishing [14], Lemmas 6.1.2 and 6.1.4, respectively. Furthermore, the local Lipschitz continuity of the quasi-potential is never used in the proof of (1.12) and (1.13), while [14], Lemma 6.1.3, can be bypassed (since it is only used for proving [14], Lemma 6.1.4). The LDP and [14], Lemmas 6.1.1–6.1.4, together imply [14], Lemmas 6.1.5–6.1.9, containing the fundamental transition times estimates for the establishment of [14], Lemmas 6.2.1, 6.2.2, proving that $\mathcal{V}_{\mathcal{D}}$ is the relevant functional for the estimation of transition probabilities between K_i's. The combination of these results finally yields (1.12) and (1.13) as explained in the proofs of [14], Theorems 6.5.1, 6.5.3. We thus proceed to state and prove the adaptations of [14], Lemmas 6.1.1 and 6.1.4, to the current setting.

LEMMA 4.1. *For $\mathcal{D} \subset \mathbb{R}_+^d$ as in Theorem 1.15, there exist $\varkappa \geq 1$, $\varepsilon > 0$ and $C(t) \to 0$ (as $t \to 0$), such that for any $x, y \in \mathcal{D}$ with $\|x - y\|_1 < \varepsilon$, there exists a path $z(\cdot) \subset \mathcal{D}$, of length $t = \varkappa\|x - y\|_1$ with $I_{x,t}(z) \leq C(t)$ and $z(t) = y$.*

PROOF. By the continuity of $\lambda_r(\cdot)$ of (1.5) on $\mathcal{D}$ compact, $\bar{\lambda} := \max_{r\in\mathcal{R},x\in\mathcal{D}}\{\lambda_r(x)\}$ is finite. Further, since $\mathrm{Co}\{c^r\}_{r\in\mathcal{R}} = \mathbb{R}^d$ the sets $Q_{\mathcal{R}}(\xi)$ are nonempty and

$$\bar{q} := e \vee \max_{\|\xi\|_1 \leq 1} \min\bigl\{\|q\|_\infty : q \in Q_{\mathcal{R}}(\xi)\bigr\} < \infty.$$

Setting $\bar{c}_\star := \sup_{r\in\mathcal{R}}\{\|c^r_{\text{in}}\|_1\}$ and $\gamma := \bar{\lambda} - \bar{q} + \bar{q}\log(\bar{q}/\min_{r\in\mathcal{R}}\{k_r \wedge 1\})$ for the reaction constants k_r of (1.5), we then have for any $z \in \mathcal{D}$ and $\|\xi\|_1 \leq 1$ the bound

$$L\bigl(\lambda(z), \xi\bigr) \leq m\Bigl[\gamma + \bar{q}\bar{c}_\star\Bigl(\log \min_{i=1}^{d}\{z_i\}\Bigr)_-\Bigr]$$

on the Lagrangian of (1.7). Thus, if $z \in AC_{0,t}(\mathcal{D})$ with $z(0) = x$ is such that $\|z'(s)\|_1 \leq 1$ and $\min_i\{z_i(s)\} \geq \beta s$, then for the rate function of (1.8),

$$I_{x,t}(z) \leq c(t) := m\int_0^t \bigl[\gamma + \bar{q}\bar{c}_\star(\log\beta s)_-\bigr]\,\mathrm{d}s. \tag{4.1}$$

Similar to [29], Lemma 2.1, Assumption A.3 implies that for some $\beta \in (0, 1)$, $\varepsilon \in (0, 1/3)$ and $v^{(j)} \in \mathbb{R}^d$ with $\|v^{(j)}\|_1 \leq 1$, there exists a finite covering of $\mathcal{D}$ by balls $\{\mathcal{B}_j\}$ such that

$$\min_{\tilde{x}\notin\mathcal{D}}\|x + sv^{(j)} - \tilde{x}\|_\infty \geq \beta s \qquad \forall x \in \mathcal{D} \cap \mathcal{B}_j^\varepsilon, s \leq \varepsilon/\beta. \tag{4.2}$$

Fixing such a covering we set $\varkappa = 1 + 2/\beta$. Suppose now that $x \in \mathcal{D} \cap \mathcal{B}_j$ and $\|y - x\|_1 = \delta < \varepsilon$ for some $y \in \mathcal{D}$. Taking $t = t_1 + t_2 + t_3$ for $t_1 = t_3 = 2\delta/\beta$ and $t_2 = \delta$, consider the continuous path from $x^{(1)} := x$ to $x^{(4)} := y$, composed of the line segments between $x^{(1)}$, $x^{(2)} = x^{(1)} + t_1 v^{(j)}$, $x^{(3)} = x^{(4)} + t_3 v^{(j)}$ and $x^{(4)}$. That is, $z^{(1)}(s) = x^{(1)} + sv^{(j)}$ for $s \in [0, t_1]$, then $z^{(2)}(s) = x^{(2)} + \frac{s}{\delta}(y - x)$ for $s \in [0, t_2]$, and finally, in reverse $z^{(3)}(s) = x^{(4)} + sv^{(j)}$ for $s \in [0, t_3]$. Since $y \in \mathcal{D} \cap \mathcal{B}_j^\delta$ and $\delta \leq \varepsilon$, it follows from (4.2) that $\min_i\{z_i^{(\ell)}(s)\} \geq \beta s$ and $z^{(\ell)}(s) \in \mathcal{D}$ for $\ell = 1, 3$ and $s \in [0, \delta/\beta]$. The end points $x^{(2)}$ and $x^{(3)}$ of $z^{(2)}(\cdot)$, are δ apart and by the preceding, of at least 2δ sup-distance from $\mathcal{D}^c$. Consequently, $\inf_{\xi\in\mathcal{D}^c}\|z^{(2)}(s) -$

$\xi\|_1 \geq \delta$ and $\min_i\{z_i^{(2)}(s)\} \geq \delta \geq \beta s$ for $s \in [0, \delta]$. By construction, $\|z'^{(\ell)}(s)\|_1 \leq 1$ for $\ell = 1, 2, 3$ and all s, so in view of (4.1)

$$I_{x,t}(z) = \sum_{\ell=1}^{3} I_{x^{(\ell)},t_\ell}(z^{(\ell)}) \leq \sum_{\ell=1}^{3} c(t_\ell) =: C(t),$$

as claimed. □

LEMMA 4.2. *Let $\mathcal{D}_{-\delta} := \mathcal{D} \setminus (\partial\mathcal{D})^\delta$ with $\mathcal{D}$ as in Assumption A.3. For some $C_\star(t) \to 0$, some $\eta(\gamma, \varkappa_\star, \mathcal{D}) > 0$, any $\varkappa_\star < \infty$, $\gamma > 0$ and $\delta \in (0, \eta)$, if $T + I_{z_0,T}(z) \leq \varkappa_\star$ for $z([0, T]) \subset \mathcal{D}$, then there exists $\widetilde{T} \leq T + 3\varkappa\gamma$ and $\tilde{z}([0, \widetilde{T}]) \subset \mathcal{D}_{-\delta}$ such that $I_{\tilde{z}_0,\widetilde{T}}(\tilde{z}) \leq I_{z_0,T}(z) + C_\star(\gamma)$ and $\|\tilde{z}(0) - z(0)\|_1 + \|\tilde{z}(\widetilde{T}) - z(T)\|_1 \leq 2\delta$. The same holds for $\mathcal{D}_{+\delta} := \mathcal{D}^\delta \cap \mathbb{R}_+^d$ and $\mathcal{D}$, instead of $\mathcal{D}$ and $\mathcal{D}_{-\delta}$, respectively.*

PROOF. From [29], Lemma 2.1, and Assumption A.3 we have [29], Assumption 2.1, holding. Further, with $\bar{\lambda}$ finite, the path $z(\cdot)$ whose length and rate function are both bounded by $\varkappa_\star$, makes at most $J = J(\varkappa_\star)$ transitions between the balls $\mathcal{B}_j$ in the covering of $\mathcal{D}$ (see [29], Lemma 3.5). Each of the monomials $\lambda_r(\cdot)$ of (1.5) is $c_\mathcal{D}$-Lipschitz continuous on the compact $\mathcal{D}$ and nondecreasing along any short path that originates in a small enough neighborhood of the set of zeroes of $\lambda_r(\cdot)$ in $\partial\mathbb{R}_+^d$, and is directed inward to $(\mathbb{R}_+^d)^o$. In particular, for some $\nu > 0$ and all j, w.l.o.g. the vectors $v^{(j)}$ in (4.2) are such that $\lambda_r(x + \alpha v^{(j)}) \geq \lambda_r(x)$ for any $\alpha \in [0, \nu]$ and $x \in \mathcal{B}_j$ for which $\lambda_r(x) \leq \nu$.

Adapting [29], Lemma 4.3, we construct for $\beta \in (0, 1)$ as in the proof of Lemma 4.1 and some $\eta(\gamma, \varkappa_\star, \mathcal{D}) > 0$, a path $\hat{z} \in (\mathcal{D})_{-2\eta}$ with $I_{\hat{z}_0,\hat{T}}(\hat{z}) \leq I_{z_0,T}(z) + 2\gamma$, $\sup_t \|\hat{z}(t) - z(t)\|_1 \leq \gamma$, $\hat{T} \leq T + \gamma$ and $\|\hat{z}_0 - z_0\|_1 \leq \eta' := 4\eta/\beta$. Specifically, let $\hat{z}_0 = z_0 + \eta' v^{(i)}$ or $\hat{z}_0 = z_0$ depending on whether $z_0 \in \mathcal{B}_i$ for $\mathcal{B}_i$ touching, or not touching, $\partial\mathcal{D}$, respectively. Thereafter, $\hat{z}(\cdot)$ is parallel to $z(\cdot)$, except that at the kth time the path $z(\cdot)$ transitions to a new ball $\mathcal{B}_j$ of the covering (that touches $\partial\mathcal{D}$), a linear segment in direction $v^{(j)}$ is inserted in $\hat{z}(\cdot)$ for duration $\eta_k = \eta'(3/\beta)^k$, to keep it within $\mathcal{D}_{-2\eta}$. With at most $J(\varkappa_\star)$ transitions between different balls $\mathcal{B}_j$, taking $\eta > 0$ small enough guarantees that the total contribution of time shifts to the length $\hat{T}$ of the path $\hat{z}$ be at most γ, and that $\sup_s \|\hat{z}(s) - z(s)\|_1 \leq \gamma$. Next, having $I_{x,t}(x + s v^{(j)}) \leq c(t)$, due to (4.1), the rate contribution of all additional linear segments is at most $\sum_k c(\eta_k) \leq \gamma$ (for small enough $\eta > 0$). Taking even smaller $\eta > 0$, bounds by γ (uniformly over all such path z), the accumulated rate difference between pieces of $\hat{z}(\cdot)$ and their parallels within $z(\cdot)$, as soon as we show that for some $g_\mathcal{D}(\alpha) \to 0$ when $\alpha \to 0$,

$$z(\cdot) \subset \mathcal{B}_j, \alpha \in [0, \nu/c_\mathcal{D}] \quad \Rightarrow \quad I_{z_0,t}(z(\cdot) + \alpha v^{(j)}) \leq I_{z_0,t}(z(\cdot)) + t g_\mathcal{D}(\alpha). \tag{4.3}$$

To this end, if $|\lambda_r - \hat{\lambda}_r| \le c_{\mathcal{D}}\alpha$ and $\hat{\lambda}_r \ge \lambda_r$ whenever $\lambda_r \le \nu$, then by (1.7), for any $\xi \in \mathbb{R}^d$,

$$L(\hat{\lambda}, \xi) - L(\lambda, \xi) \le \|\hat{\lambda} - \lambda\|_1 + \max_r \left\{ \log\left(\frac{\hat{\lambda}_r}{\lambda_r}\right)\right\}_- \le m c_{\mathcal{D}}\alpha - \log(1 - c_{\mathcal{D}}\alpha/\nu),$$

hence denoting the RHS by $g_{\mathcal{D}}(\alpha)$ yields (4.3) [see (1.8)].

Now, fixing $\delta \in (0, \eta)$, let $\tilde{z}(\cdot)$ be $\hat{z}(\cdot)$ augmented by the initial/final piecewise linear path of Lemma 4.1, leading from $\tilde{z}(0) := \arg\min_{z \in \mathcal{D}_{-\delta}} \|z - z(0)\|_1$ to $\hat{z}(0)$ and from $\hat{z}(\hat{T})$ to $\tilde{z}(\tilde{T}) := \arg\min_{z \in \mathcal{D}_{-\delta}} \|z - z(T)\|_1$, respectively. Since both $\|\tilde{z}(0) - \hat{z}(0)\|_1 \le \delta + \gamma$ and $\|\hat{z}(\hat{T}) - \tilde{z}(\tilde{T})\|_1 \le 2\gamma + \delta$, taking $\eta \le \gamma \le \varepsilon/3$ we have by Lemma 4.1 that the length of each augmented path is at most $\varkappa\gamma$ and its contribution to the total rate does not exceed $C(3\gamma)$. Finally, note that by construction both end-points of these initial and final pieces are in $\mathcal{D}_{-\delta}$, whereby the construction of Lemma 4.1 guarantees that their minimal distance from $\partial\mathcal{D}$ be attained at one of their end points, hence do not exceed δ. □

While (1.12) and (1.13) involve only the process $t \mapsto X_t^v$ within the compact $\mathcal{D}$, this is not the case for (1.14) which is established in [14], Theorem 6.6.2, under the additional assumption of a compact state space, which we lack here. However, the latter proof applies for the stopping time $\tau_{\pi,\varrho} := \tau_\pi \wedge \sigma_\varrho$ and the nonrandom $C_\varrho(\pi)$ obtained via [14], equations (6.6.1), (6.6.2), from $I_{x,t}^{(\varrho)}(\cdot)$ of (1.8) that corresponds to $\lambda_r(x)\mathbb{I}_{\widetilde{K}_\varrho}(x)$, with $\lambda_r(\cdot)$ of (1.5) and $\widetilde{K}_\varrho$ of (2.2) [as the Markov jump processes $X_t^{v,\varrho}$ from the proof of Theorem 1.6 are $\widetilde{K}_\varrho$-valued and satisfy the LDP with rate functions $I_{x,t}^{(\varrho)}(\cdot)$]. For $\varrho \ge \gamma$ and $\bigcup_j K_j^\delta \subset \widetilde{K}_\gamma$ it is easy to verify that using $I_{x,t}^{(\varrho)}(\cdot)$ instead of $I_{x,t}(\cdot)$ amounts to replacing the quasi-potential $\mathcal{V}(\cdot,\cdot)$ by $\mathcal{V}_{\widetilde{K}_\varrho}(\cdot,\cdot)$, with an additional attractor of the dynamics at $(\widetilde{K}_\varrho)^c$. It is irrelevant that Assumption A.4 fails for this new attractor, since it is outside π hence the transitions $(\widetilde{K}_\varrho)^c \to K_j$ play no role in $C_\varrho(\pi)$. By the same reasoning, the rate $I_{x,t}(z)$ of any path $z(\cdot)$ exiting $\widetilde{K}_\varrho$ is part of the minimization yielding $C_\varrho(\pi)$, while those paths which are confined to $\widetilde{K}_\varrho$ make exactly the same contribution to $C_\varrho(\pi)$ and to $C(\pi)$. Consequently, $C_\varrho(\pi) \uparrow C_\infty(\pi) \le C(\pi)$ and $v^{-1}\log\tau_{\pi,\varrho} \to C_\infty(\pi)$ when $v \to \infty$ followed by $\varrho \to \infty$. The compact sets $\widetilde{K}_\varrho$ satisfy Assumption A.3, so by Lemma 4.1 the quasi-potential $\mathcal{V}(x, y)$ is everywhere finite. This implies that $C(\pi)$ is finite, and thereby so is $C_\infty(\pi)$. Considering Lemma 2.1 for some $\beta > C_\infty(\pi)$ and $\varrho \to \infty$, we thus conclude that $v^{-1}\log\tau_\pi \to C_\infty(\pi)$, which translates to (1.14) provided $C_\infty(\pi) \ge C(\pi)$. The latter is a direct consequence of our next lemma, showing that $\mathcal{V}(\widetilde{K}_\gamma, (\widetilde{K}_\varrho)^c) \to \infty$ as $\varrho \to \infty$. Indeed, the second term on the RHS of [14], equation (6.6.2), is independent of the addition of $(\widetilde{K}_\varrho)^c$ to the set of attractors [hence identical for $C(\pi)$ and $C_\varrho(\pi)$], while every element over which the minimum is taken in [14], equation (6.6.1), is either the same for $C(\pi)$ and $C_\varrho(\pi)$, or involves some transition $K_j \to (\widetilde{K}_\varrho)^c$. Since $\mathcal{V}(\cdot,\cdot) \geq 0$, terms involving any such transition are irrelevant when $\mathcal{V}(\widetilde{K}_\gamma, (\widetilde{K}_\varrho)^c) > C(\pi)$.

LEMMA 4.3. *Under Assumption A.1, for any γ finite,*

$$\lim_{\varrho\to\infty} \inf_{t\geq 0}\{J_\gamma(t,\varrho)\} = \infty, \tag{4.4}$$
$$J_\gamma(t,\varrho) := \inf_{\|x\|_1\leq\gamma} \inf_{\{z(\cdot):\sup_{s\leq t}\|z(s)\|_1>\varrho\}} \{I_{x,t}(z)\}.$$

PROOF. The lower bound of the LDP of Theorem 1.6 for the open set $\Gamma := \{z : z(t) \in (\widetilde{K}_\varrho)^c \text{ for some } t \leq T\}$, implies that

$$-J_\gamma(T,\varrho) \leq \liminf_{v\to\infty} \frac{1}{v} \log\Bigl(\sup_{\|x_0^v\|_1\leq\gamma} \mathbb{P}_{x_0^v}\Bigl[\sup_{t\in[0,T]} \|X_t^v\|_1 > \varrho\Bigr]\Bigr). \tag{4.5}$$

While proving Lemma 2.1, we saw that the RHS of (4.5) is, for some finite $\varkappa = \varkappa(\gamma)$, with the constant b of Assumption A.1(a), any T and $\varrho \geq \varrho(\ell)$, at most

$$\limsup_{v\to\infty} v^{-1} \log\{\ell^{-v}[e^{\varkappa v} + Te^{bv}]\} = -\log\ell + \varkappa \vee b. \tag{4.6}$$

Combining (4.5) and (4.6), we establish (4.4) upon taking $\varrho \to \infty$ followed by $\ell \to \infty$. □

Acknowledgments. We thank Ofer Zeitouni for pointing our attention to [9]. Furthermore, AA and JPE would like to thank Noé Cuneo and Neil Dobbs for helpful discussions.

REFERENCES

[1] ANDERSON, D. F. (2011). A proof of the global attractor conjecture in the single linkage class case. *SIAM J. Appl. Math.* **71** 1487–1508. MR2835244

[2] ANDERSON, D. F., CRACIUN, G., GOPALKRISHNAN, M. and WIUF, C. (2015). Lyapunov functions, stationary distributions, and non-equilibrium potential for reaction networks. *Bull. Math. Biol.* **77** 1744–1767. MR3423060

[3] BALÁZSI, G., VAN OUDENAARDEN, A. and COLLINS, J. J. (2011). Cellular decision making and biological noise: From microbes to mammals. *Cell* **144** 910–925.

[4] BISHOP, L. M. and QIAN, H. (2010). Stochastic bistability and bifurcation in a mesoscopic signaling system with autocatalytic kinase. *Biophys. J.* **98** 1–11.

[5] BUDHIRAJA, A., DUPUIS, P., FISCHER, M. and RAMANAN, K. (2015). Local stability of Kolmogorov forward equations for finite state nonlinear Markov processes. *Electron. J. Probab.* **20** no. 81, 30. MR3383565

[6] CAPPELLETTI, D. and WIUF, C. (2016). Product-form Poisson-like distributions and complex balanced reaction systems. *SIAM J. Appl. Math.* **76** 411–432. MR3463706

[7] CRACIUN, G. (2015). Toric differential inclusions and a proof of the global attractor conjecture. ArXiv E-prints arXiv:1501.02860.

[8] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics* (*New York*) **38**. Springer, New York. MR1619036

[9] DUPUIS, P., RAMANAN, K. and WU, W. (2016). Large deviation principle for finite-state mean field interacting particle systems. ArXiv Preprint arXiv:1601.06219.

[10] DYNKIN, E. B. (1965). *Markov processes*. *Vols*. *I*, *II*. *Die Grundlehren der Mathematischen Wissenschaften* **121, 122**. Academic Press, New York; Springer, Berlin. Translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra, G. Majone. MR0193671

[11] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes*: *Characterization and Convergence*. Wiley, New York. MR838085

[12] FEINBERG, M. (1987). Chemical reaction network structure and the stability of complex isothermal reactors—I. The deficiency zero and deficiency one theorems. *Chem*. *Eng*. *Sci*. **42** 2229–2268.

[13] FENG, J. and KURTZ, T. G. (2006). *Large Deviations for Stochastic Processes*. *Mathematical Surveys and Monographs* **131**. Amer. Math. Soc., Providence, RI. MR2260560

[14] FREIDLIN, M. I. and WENTZELL, A. D. (1984). *Random Perturbations of Dynamical Systems*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **260**. Springer, New York. MR0722136

[15] GAVEAU, B., MOREAU, M. and TOTH, J. (1999). Variational nonequilibrium thermodynamics of reaction-diffusion systems. I. The information potential. *J*. *Chem*. *Phys*. **111** 7736–7747.

[16] GE, H. and QIAN, H. (2010). Non-equilibrium phase transition in mesoscopic biochemical systems: From stochastic to nonlinear dynamics and beyond. *J*. *R*. *Soc*. *Interface* **8** 107–116.

[17] GILLESPIE, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J*. *Phys*. *Chem*. **81** 2340–2361.

[18] GOPALKRISHNAN, M., MILLER, E. and SHIU, A. (2014). A geometric approach to the global attractor conjecture. *SIAM J*. *Appl*. *Dyn*. *Syst*. **13** 758–797.

[19] GUPTA, A. and KHAMMASH, M. (2014). Determining the long-term behavior of cell populations: A new procedure for detecting ergodicity in large stochastic reaction networks. *IFAC Proceedings Volumes* **47** 1711–1716.

[20] HORN, F. and JACKSON, R. (1972). General mass action kinetics. *Arch*. *Ration*. *Mech*. *Anal*. **47** 81–116.

[21] JOHNSTON, M. D., PANTEA, C. and DONNELL, P. (2016). A computational approach to persistence, permanence, and endotacticity of biochemical reaction systems. *J*. *Math*. *Biol*. **72** 467–498.

[22] LI, T. and LIN, F. (2016). Two-scale large deviations for chemical reaction kinetics through second quantization path integral. *J*. *Phys*. *A*: *Math*. *Theor*. **49** 135204.

[23] MILLER, E. and PAK, I. (2008). Metric combinatorics of convex polyhedra: Cut loci and nonoverlapping unfoldings. *Discrete Comput*. *Geom*. **39** 339–388. MR2383765

[24] NOBILE, M. S., CIPOLLA, D., CAZZANIGA, P. and BESOZZI, D. (2016). GPU-powered evolutionary design of mass-action-based models of gene regulation. In *Evolutionary Computation in Gene Regulatory Network Research* 118–150. Wiley, New York.

[25] PAULEVÉ, L., CRACIUN, G. and KOEPPL, H. (2014). Dynamical properties of discrete reaction networks. *J*. *Math*. *Biol*. **69** 55–72.

[26] REY-BELLET, L. (2006). Ergodic properties of Markov processes. In *Open Quantum Systems II*: *The Markovian Approach*. *Lecture Notes in Math*. **1881** 1–39. Springer, Berlin. MR2248986

[27] ROCK, K., BRAND, S., MOIR, J. and KEELING, M. J. (2014). Dynamics of infectious diseases. *Rep*. *Progr*. *Phys*. **77** 026602, 51. MR3159087

[28] SHWARTZ, A. and WEISS, A. (1995). *Large Deviations for Performance Analysis*: *Queues*, *Communications*, *and Computing*. Chapman & Hall, London. MR1335456

[29] SHWARTZ, A. and WEISS, A. (2005). Large deviations with diminishing rates. *Math. Oper. Res.* **30** 281–310. MR2142034

A. AGAZZI
DEPARTMENT OF MATHEMATICS
DUKE UNIVERSITY
120 SCIENCE DRIVE
RM 117 PHYSICS BLDG
DURHAM, NORTH CAROLINA 27708
USA
E-MAIL: agazzian@gmail.com

A. DEMBO
DEPARTMENT OF MATHEMATICS AND STATISTICS
STANFORD UNIVERSITY
390 SERRA MALL
STANFORD, CALIFORNIA 94305
USA
E-MAIL: amir@math.stanford.edu

J.-P. ECKMANN
DÉPARTEMENT DE PHYSIQUE THÉORIQUE
AND SECTION DE MATHÉMATIQUES
UNIVERSITÉ DE GENÈVE
QUAI ERNEST ANSERMET 24
1211 GENEVA, CH
SWITZERLAND
E-MAIL: Jean-Pierre.Eckmann@unige.ch